December 1991 Final

Standardization of End-to-End Performance of Digital Video Teleconferencing/Video Telephony Systems

C-DCA100-91-C-0031

Delta Information Systems, Inc.
300 Welsh Road, Suite 120
Horsham, PA 19044-2273

National Communications System
Office of Technology & Standards
701 S. Court House Road
Arlington, VA 22204-2198

NCS TIB 91-15

Approved for public release; distribution unlimited.

This report covers the standardization of end-to-end performance for full motion videophone and video teleconferencing. It investigates the criteria for such a standardization on an end-to-end basis to ensure delivery of a specified quality of service for various Governement TV applictions. It is distinguished from the previous effort covered in TIB 90-6 by concentrating on a codec designed in accordance with CCITT standard H.261. Section 2 of this report describes the available background material consisting of both previous NCS programs and related efforts performed in conjunction with the deliberations of ANSI Committee T1Q1.5. Section 3 contains the various criteria that had to be considered, concentrating on the test material and the characteristics of the codec under test. Section 4 describes the subjective evaluation of the processed test tapes which were rated in terms of picture impairment by a panel of impartial evaluators. Section 5 covers objective tests and includes the overall methodology and test results for three parameters which have previously been identified as important for the objective evaluation of codec performance. Correlation between subjective and objective test results is examined in Section 6. Section 7 discusses the overall results and adds a few not (cont'd)

Video teleconferencing 100
Full motion videophone

Unclassified Unclassified Unclassified Unlimited

Item #13 (Abstract) Cont'd:

previously mentioned factors. Section 8 briefly summarizes the total program and suggests a number of follow-on tests which together will form a solid basis for an end-to-end performance standard for digital video teleconferencing/video telephony systems.

NCS TIB 91-15

NATIONAL COMMUNICATIONS SYSTEM

TECHNICAL INFORMATION BULLETIN 91-15

STANDARDIZATION OF END-TO-END PERFORMANCE OF DIGITAL VIDEO TELECONFERENCING/VIDEO TELEPHONY SYSTEMS

DECEMBER 1991

OFFICE OF THE MANAGER
NATIONAL COMMUNICATIONS SYSTEM

ARLINGTON, VA 22204-2199

NCS TECHNICAL INFORMATION BULLETIN 91-15

STANDARDIZATION OF END-TO-END
PERFORMANCE OF DIGITAL VIDEO
TELECONFERENCING/VIDEO TELEPHONY SYSTEMS

PROJECT OFFICER	APPROVED FOR PUBLICATION:
	Dennis Bodson
GARY REKSTAD Electronics Engineer Office of Technology and Standards	DENNIS BODSON Assistant Manager Office of Technology and Standards

FOREWORD

Among the responsibilities assigned to the Office of the Manager, National Communications System, is the management of the Federal Telecommunication Standards Program. Under this program, the NCS, with the assistance of the Federal Telecommunication Standards Committee identifies, develops, and coordinates proposed Federal Standards which either contribute to the interoperability of functionally similar Federal telecommunication systems or to the achievement of a compatible and efficient interface between computer and telecommunication systems. In developing and coordinating these standards, a considerable amount of effort is expended in initiating and pursuing joint standards development efforts with appropriate technical committees of the International Organization for Standardization, and the International Telegraph and Telephone Consultative Committee of the International Telecommunication Union. This Technical Information Bulletin presents and overview of an effort which is contributing to the development of compatible Federal, national, and international standards in the area of Video Teleconferencing Quality. It has been prepared to inform interested Federal activities of the progress of these efforts. Any comments, inputs or statements of requirements which could assist in the advancement of this work are welcome and should be addressed to:

Office of the Manager
National Communications System
Attn: NT
701 S. Court House Road
Arlington, VA 22204-2199

STANDARDIZATION OF END-TO-END PERFORMANCE OF DIGITAL VIDEO TELECONFERENCING/VIDEO TELEPHONY SYSTEMS

December, 1991

FINAL REPORT
DCA100-91-C-0031
TASK ORDER NO. 91-004

Submitted to:
NATIONAL COMMUNICATIONS SYSTEM
WASHINGTON, DC

DELTA INFORMATION SYSTEMS, INC.
300 Welsh Road, Ste. 120
Horsham, PA 19044-2273
TEL: (215) 657-5270 FAX: (215) 657-5273

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION AND SUMMARY

This document summarizes work performed by Delta Information Systems, Inc. (DIS) for the National Communications System (NCS), Office of Technology and Standards. This office is responsible for the management of the Federal Telecommunications Standards Program, which develops telecommunications standards, whose use is mandatory for all Federal departments and agencies. This study was performed under task order number 91-004 of contract number DCA100-91-C-0031.

This report covers the standardization of end-to-end performance for full motion videophone and video teleconferencing. It investigates the criteria for such a standardization on an end-to-end basis to ensure delivery of a specified quality of service for various Government TV applications. It is distinguished from the previous effort covered in NCS-TIB-90-6 by concentrating on a codec designed in accordance with CCITT Standard H.261. Even though most US Government applications are domestic and as such could base its standard on any equipment using a proprietary codec algorithm, the general aspects of digital video communication are more and more of a global nature. It is advantageous for the US Government to stay in step and be compatible with the rest of the world. Furthermore, all US and foreign codec manufacturers will obviously have to concentrate their future design and improvement efforts on the world-wide commercial market. This means that all future codecs with upgraded performance will incorporate the H.261 algorithm while proprietary algorithms are becoming obsolescent even at this time.

Section 2.0 of this report describes the available background material consisting of both previous NCS programs and related efforts performed in conjunction with the deliberations of ANSI Committee T1Q1.5. Section 3.0 contains the various criteria that had to be considered, concentrating on the test material and the characteristics of the codec under test. Test tapes containing all required test material were taken to the lab of a local TV system design contractor who could make the codec available. Section 4.0 describes the subjective evaluation of the processed test tapes which were rated in terms of picture impairment by a panel of impartial evaluators. The analysis of the individual test scores and the overall performance rating at each bit rate and operating mode tested are included. Section 5.0 covers objective tests and includes the overall methodology and test results for three parameters which have previously been identified as important for the objective evaluation of codec performance. Correlation between subjective and objective test results is examined in Section 6.0. Section 7.0 discusses the overall results and adds a few not previously mentioned factors. Section 8.0 briefly summarizes the total program and suggests a number of follow-on tests which together will form a solid basis for an end-to-end performance standard for digital video teleconferencing/video telephony systems.

2.0 BACKGROUND

2.1 Previous NCS Programs

Three previous programs cover efforts which are very similar to the tasks required for the implementation of this program. They are: Standardization of End-to-End Performance for Full Motion Video Teleconferencing, NCS-TIB-90-6, Temporal Frequency as a Technique to Measure the Ability of a Teleconferencing System to Reproduce Motion, NCS-TIB-91-2, and Subjective Tests of Teleconferencing Codecs, NSC-TIB-91-6. Many of the test methodologies used in this program are identical to the ones developed for these previous programs and are being incorporated in this report only in abbreviated form.

2.2 T1Q1.5 Committee Contributions

Delta's continuing very active participation in the efforts of the Subworking Group on Video Teleconferencing/Video Telephony of the American National Standards Institute (ANSI). Committee T1Q1.5 produced a number of valuable contributions pertaining to both subjective and objective testing of video codecs. Many experts from both research and operating organizations have furnished useful material which eventually will provide the basis for a specification encompassing performance requirements for typical video codec applications test parameters and practical test methods. The deliberations of this group are very detailed and a completed standard should not be expected in the near future. The latest draft version of the standard is attached as Appendix A.

3.0 TEST CRITERIA

3.1 Codec Characteristics

Testing of analog video systems uses well established parameters and performance limits. This is not the case with digital video codecs which require additional test parameters and have different performance requirements depending on their application. Since the field is still quite new, a universally applicable testing baseline has not yet been established. Codec testing requirements can be strongly influenced by the codec design.

Even though the tests described in the final reports referenced in Section 2.0 were performed on three distinctly different codecs of proprietary design, the equipments, though not interoperable, are fundamentally sufficiently similar to produce comparable results. On the other hand, CCITT standard H.261 has the unique purpose to provide interoperability between different video standards and equipments from different manufacturers. This makes the basic algorithm and its directly associated performance features quite firm but allows considerable freedom in implementation of required functions, incorporating optional features and new developments for improved performance, and making trade-offs between performance parameters. This large amount of flexibility produces a wide range of possible performance features which must be accommodated by the test methodology.

Figure 3.1 shows a very simplified block diagram of an H.261 codec in an

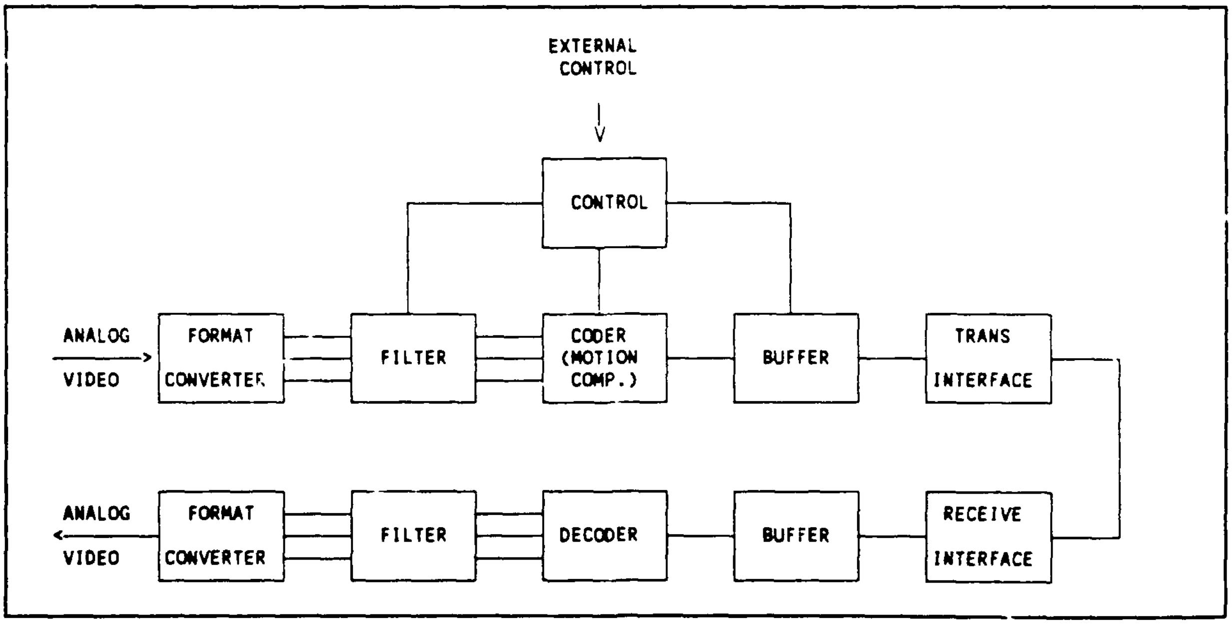


FIGURE 3.1
H.261 CODEC VIDEO TRANSMISSION SYSTEM

end-to-end video transmission system including both firmly specified and peripheral flexible functions. The format converter changes either NTSC or PAL inputs to a common intermediate format (CIF) featuring sequential scanning at a nominal frame rate of 30 per second. It also breaks down the composite color video into components, one luminance and two color difference signals. In the full CIF format, the luminance component is sampled in 288 lines and 352 samples per line while the color difference components use 144 lines with 176 samples per line. The lower quality Quarter CIF (QCIF) format has half the number of lines and samples.

Each of the component signals passes through a shaping filter as required by the subsequent sampling process to prevent aliasing. A filter characteristic applying to the same sampling rate as used for the CIF luminance signal is given in CCIR Rec. 601-1 as shown on Figure 3.2. This characteristic is not specified in H.261 but is only typical, the details producing best performance are up to the ingenuity of the equipment designer. The frequency values are halved for the QCIF format and halved again for the color difference signals. This filter determines the achievable limits of static (or spatial) picture resolution.

The video coder performs the main functions of the H.261 codec. It implements the standardized block structure and other elements of the algorithm which is necessary for interoperability but many details are left to the equipment designer. Motion compensation is highly desirable but not mandatory in the encoder. Similarly, frame subsampling is recommended but need not be incorporated. This feature results in the non-transmission of one or several frames between transmitted ones. It restricts the maximum picture bit rate by limiting the transmitted frame rate (TFR). This feature is mainly controlled by the transmission buffer which signals the coding control to prevent either overflow or underflow. The encoding strategy to limit the required transmission bit rate to a given value is up to the equipment designer. It generally is based on a trade-off between reduced picture fidelity caused by limited quantization accuracy and reduced TFR. A minimum acceptable TFR may be selected by external control. Typical examples of other performance parameters to be set externally are the operating format (CIF or QCIF) and the value of P in the P x 64 transmission bit rate.

This large number of optional and flexible features results in a range of codec performance which manifests itself in both subjective evaluations and objective tests. The relative importance of performance parameters is likely to change with different designs and applications.

3.2 Test Material

Both objective and subjective tests were implemented with previously produced test tapes. The development of these tapes was described previously in NCS-TIB-89-2 and NCS-TIB-90-7. Only the appropriate portions of the available material were used. A brief description of these portions is included here for completeness.

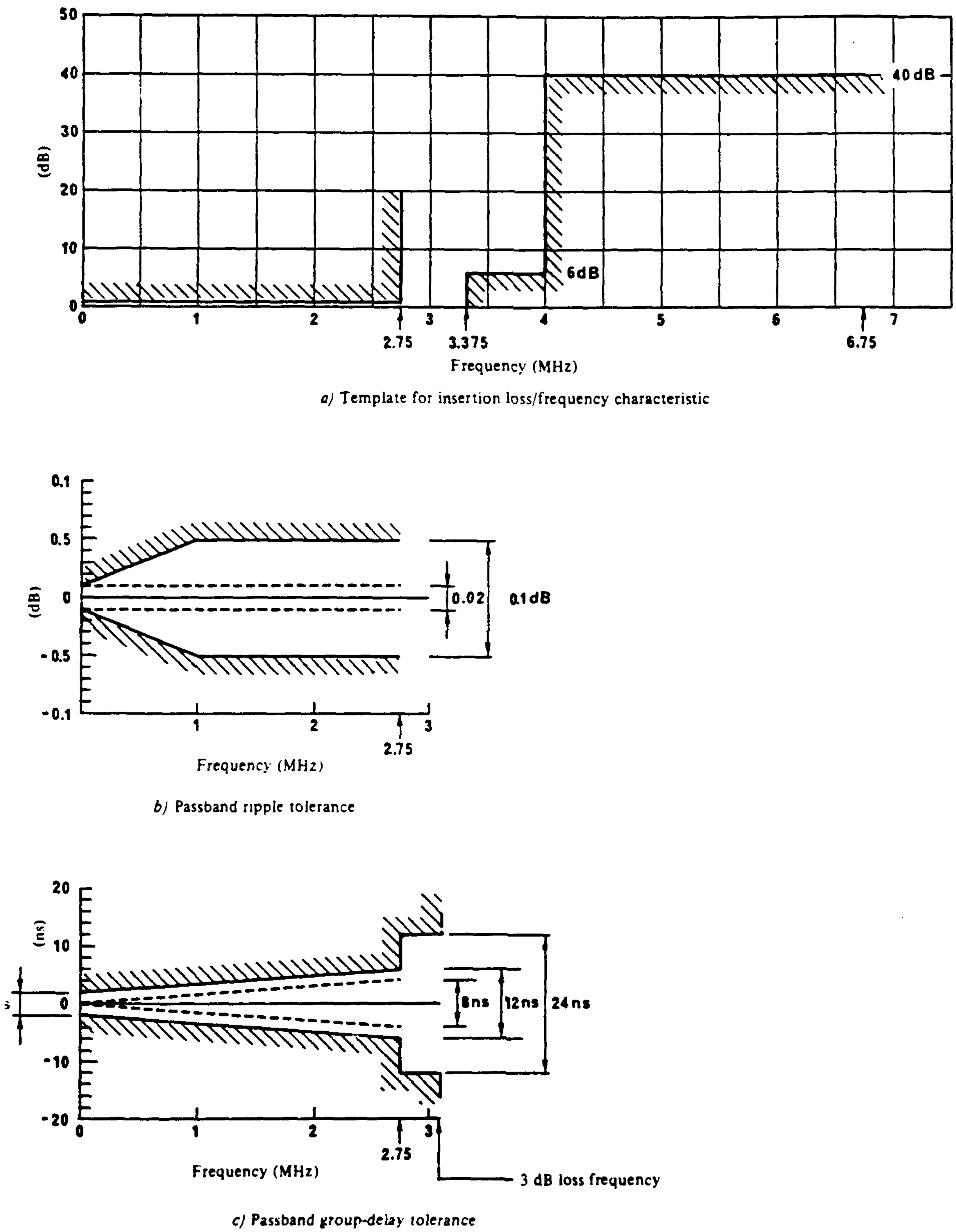


FIGURE 3.2 – *Specification for a colour-difference signal filter used when sampling at 6.75 MHz*

Note. – The lowest indicated values in *b)* and *c)* are for 1 kHz (instead of 0 MHz).

3.2.1 Subjective Test Material

Since only one H.261 codec model was available for evaluation, no differences in analog performance which could affect the correlation between subjective and objective tests can be expected. A very noticeable difference in static resolution between CIF and QCIF is inherent and predictable, therefore, static frequency response need not be measured. Previous tests have shown all other conventional analog performance parameters to be largely independent of bit rate, and only slightly changed between different codec models.

The available test material is subdivided into four parts, as follows:

Part A: Still Graphics
Part B: Motion Graphics
Part C: Limited Motion
Part D: Full Motion

These tapes were reviewed to determine their applicability considering the performance characteristics and potential use of the codec under test. Still graphics are primarily useful for testing static resolution and color rendition both of which can also be evaluated with portions of the motion graphics tape. Furthermore, differences in static resolution are produced only by the use of CIF or QCIF and thus are readily predictable. Color rendition is mainly determined by well developed and non-critical circuits which are independent of digital processing and bit rate. The "limited motion" tape contains several scenes including also full motion which is needed for the two highest bit rates tested. Furthermore, using the same test tapes for all bit rates under test makes correlation between subjective and objective tests more meaningful.

The scenarios of the two tapes used for the tests are shown in Tables 3-1 and 3-2. The sequences were chosen to be representative of most potential codec applications. If and when desired, each sequence can readily be assigned to one or possibly several user categories and separate scores for each category can be computed.

3.2.2 Objective Test Material

Objective testing is implemented with previously developed computer generated test tapes. The rotating wheel pattern described in NCS-TIB-90-7 was used for most tests. For completeness and ease of reference, the basic parameters of all patterns are repeated on Table 3-3, while a sample of one pattern is shown on Figure 3.3. The switched dot pattern described in full detail in NCS-TIB-90-6 had only limited use. The details of all patterns are given on Table 3-4, but only the slowest switching rate of 120 frames (4 seconds) was used for the measurement of image update time.

TABLE 3 - 1

VIDEO CODEC TEST TAPE PART B: *MOTION GRAPHICS*

TAPE SCENARIO

			CONTENTS			MOTION			
Sequ. No.	Source	Subject Matter	Diagram	Pictorial	Material Moving	Zoom	Pointing	Drawing	Duration (Sec.)
1	CCV	Circuit Diagram	X			X	X		35
2	CCV	Wine Bottles & Cockpit Panel		X			X		30
3	ISACOMM	Commerce Bank Poster		X	X	X			20
4	CCV	Board Layout	X			X	X	X	31
5	CCV	Activity Diagram	X			X	X	X	30
6	ISACOMM	Grand House Poster		X	X	X			27
7	CCV	Map	X		X		X	X	35
8	CCV	Speech Diagram	X		X		X	X	40
9	ISACOMM	Phone Instrument Poster		X	X	X	X		15
10	CCV	Circuit Diagram Corrections	X				X	X	30
11	CCV	Entropy Curve	X					X	38
12	ISACOMM	Drawing on Pad	X		X			X	40
13	CCV	Inspection Report Forms	X			X	X	X	38
14	CCV	Data Transmission Poster		X		X			20
15	ISACOMM	Bar Graph	X		X	X			14
16	CCV	Organization Chart	X		X	X	X		35

TABLE 3 - 2

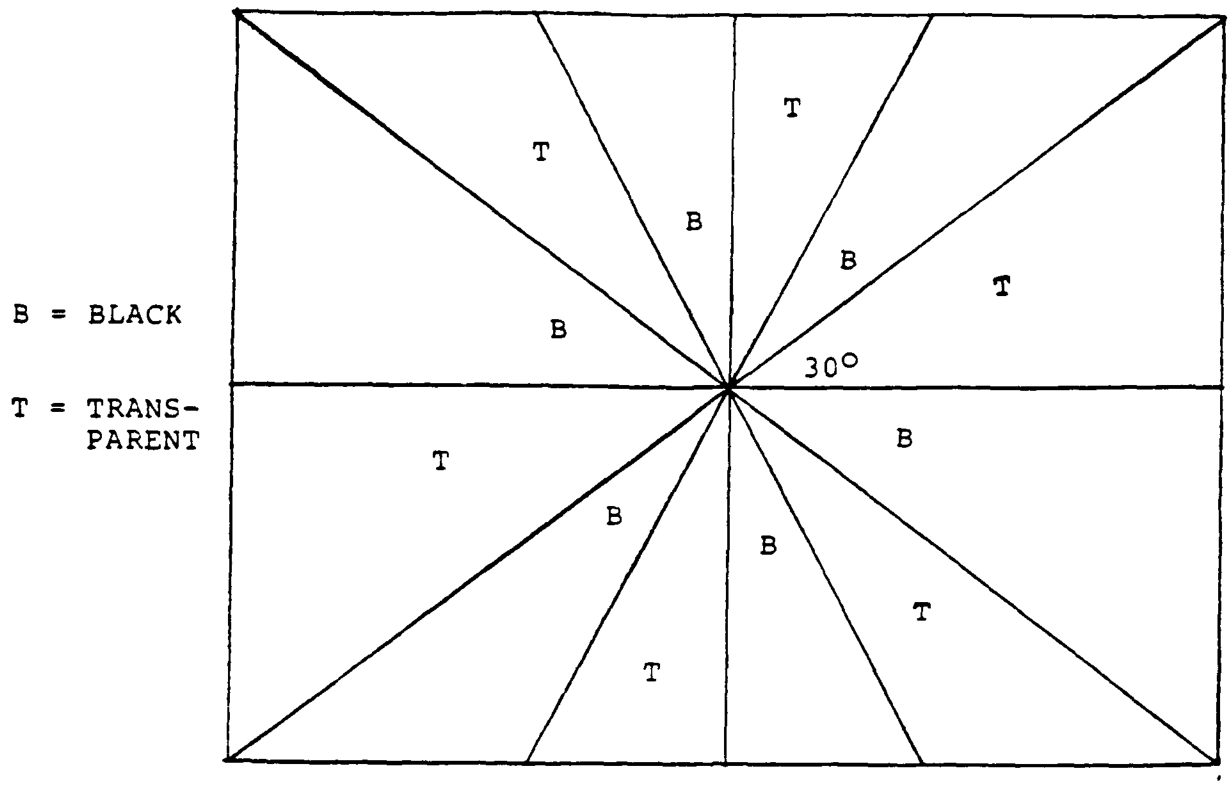
VIDEO CODEC TEST TAPE PART C: LIMITED MOTION

TAPE SCENARIO

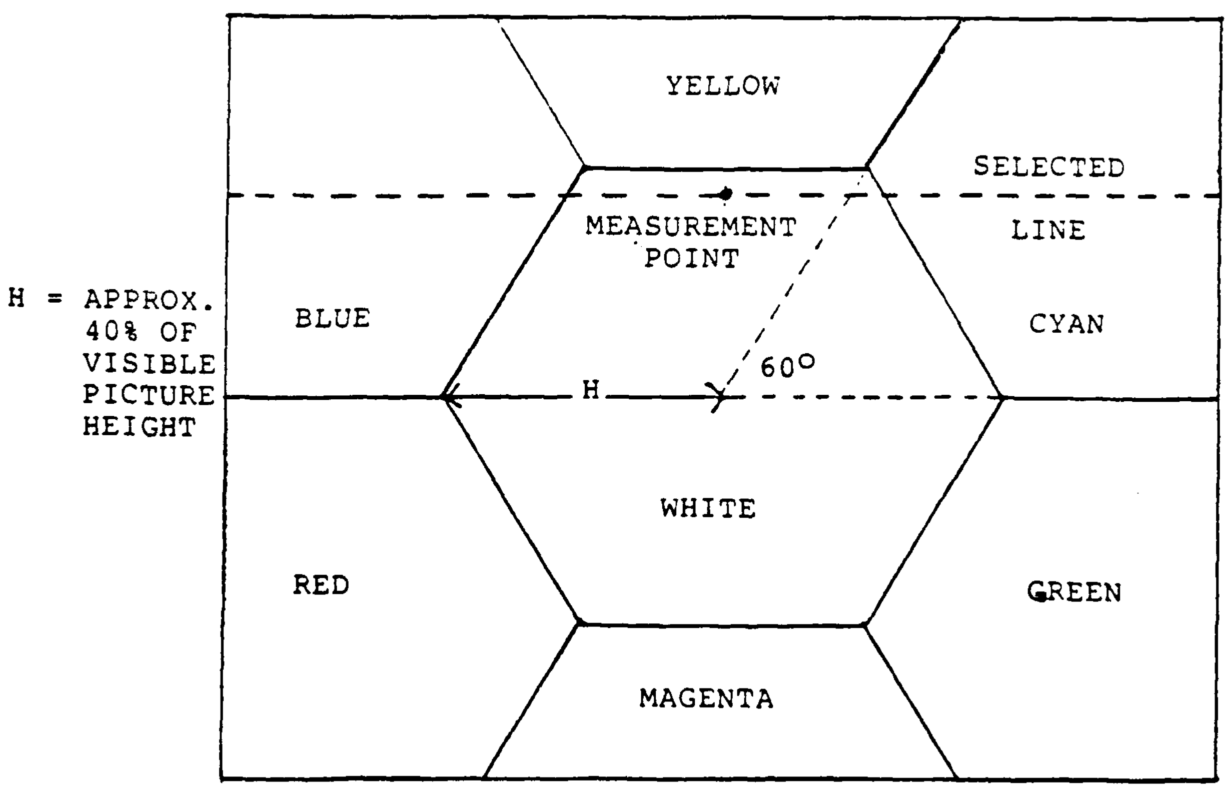
Sequ. No.	Source	Subject Matter	No. of People	Equ.	Material Phone	Handled Papers	Marker	Camera Pan	Motion Zoom	Duration (Sec.)
1	CCV	Announcer	1							20
2	CCV	Equipment Demo at Desk	1	X						51
3	CCV	"Miss America"	1							15
4	CCV	Phone Call	1		X	X				55
5	ISACOMM	Introduction of 6 People	6 X 1					X		25
6	CCV	Magazine Article	1			X				51
7	CCV	Phone Call	1			X				12
8	CCV	Equipment Demo	1	X	X					55
9	ISACOMM	Girl at Marker Board	1				X			30
10	CCV	Equipment Demo at Desk	1	X		X				80
11	CCV	Printed Board Demo	1	X				X	X	30
12	CCV	2 Groups of 3 People	2 x 3			X				30
13	CCV	Man Sitting	1							15
14	PMS	3 Groups of 2 People	3 x 2			X				43
15	SBS	2 Groups of 3 People	2 X 3							30
16	ISACOMM	6 People	6		X					15

TABLE 3-3
ROTATING WHEEL PATTERNS

TEST PATTERN NO.	SPOKE WIDTH (DEGREES)	ROTATION SPEED		TEMPORAL FREQUENCY (CPS)	FRAMES/ SPOKE	% PIXEL CHANGE/ FRAME	% BLOCK CHANGE/ FRAME
		FRAMES/ REVOLUTION	DEGREES/ SECOND				
1	30	540	20	0.33	45	2.2	18
2	30	360	30	0.50	30	3.3	
3	30	240	45	0.75	20	5.0	
4	30	180	60	1.00	15	6.7	22
5	30	144	75	1.25	12	8.3	
6	30	120	90	1.50	10	10.0	
7	30	90	120	2.00	7.5	13.3	34
8	30	72	150	2.50	6	16.7	
9	30	60	180	3.00	5	20.0	43
10	18	720	15	0.42	36	2.8	31
11	18	540	20	0.55	27	3.7	
12	18	360	30	0.85	18	5.6	
13	18	240	45	1.25	12	8.3	
14	18	180	60	1.67	9	11.1	42
15	18	144	75	2.10	7.2	13.9	
16	18	120	90	2.50	6	16.7	
17	18	90	120	3.33	4.5	22.2	57
18	10	720	15	0.75	20	5.0	50
19	10	540	20	1.00	15	6.7	54
20	10	360	30	1.50	10	10	
21	10	240	45	2.25	6.7	15	
22	10	180	60	3.00	5	20	70
23	10	144	75	3.75	4	25	75



ROTATING OVERLAY



BACKGROUND

FIGURE 3.3

ROTATING WHEEL TEST PATTERN - 30 DEGREE SPOKES

TABLE 3-4
SCENE CUT PATTERNS

CIRCLE SPACING (%)		7		4		2.25	
CIRCLE RADIUS (%)		3.25	2.25	1.75	1.25	1	.75
SWITCHING RATE (FRAMES)	TEMPORAL FREQUENCY (CPS)	TEST PATTERN NO.					
120	.125	A-1	-	A-13	-	A-25	-
60	.25	A-2	-	A-14	-	A-26	-
30	.5	A-3	-	A-15	-	A-27	-
15	1.0	A-4	-	A-16	-	A-28	-
8	1.875	A-5	A-9	A-17	A-21	A-29	A-33
4	3.75	A-6	A-10	A-18	A-22	A-30	A-34
2	7.5	A-7	A-11	A-19	A-23	A-31	A-35
1	15	A-8	A-12	A-20	A-24	A-32	A-36

NOTE: CIRCLE SPACING AND RADIUS VALUES ARE GIVEN AS PERCENTAGE OF PICTURE WIDTH.

3.3 Codec Under Test

The single P x 64 codec model available for testing operated at CIF and QCIF with values of P from 1 to 30, resulting in bit rates from 64 to 1920 Mbps. However, this particular model requires 56 Kbps for audio and an additional 2 Kbps for overhead, so that the rate available for video transmission is 58 Kbps below the nominal value. This makes 64 Kbps practically unusable. Since QCIF inherently produces a considerably degraded picture, it is likely to be used only in cases of severe bit rate restriction. Therefore, QCIF tests were performed only at 128 and 192 Kbps while CIF tests covered the whole range applicable in the US, using 128, 192, 384, 768, and 1536 Kbps.

4.0 SUBJECTIVE TESTS

4.1 Implementation

The subjective tests were implemented using CCIR Rec. 500-3 as a guide. Due consideration must be given to the fact that all test methods recommended by the CCIR apply to high quality broadcast TV systems. Digital teleconferencing codecs inherently produce much lower quality pictures which changes the emphasis on some of the details of the test methodology. The most important difference is that motion performance is one of the most significant factors in assessing the quality of a digital teleconferencing picture.

After a large series of tests performed in many countries, a method was agreed to by the European Broadcast Union (EBU) mainly for the assessment of high quality digital television pictures. This very popular method was also used in elaborate tests initiated by ANSI Committee T1Y1.1 to evaluate various algorithms for the transmission of digital TV at DS3 rates. The most significant difference in the procedure used for the tests described herein is that the impairment scale was used because it is more descriptive than the quality scale for rating the lower quality pictures produced by teleconferencing codecs. The unimpaired reference picture is shown for each test sequence ahead of the picture which was processed through the codec.

Evaluation of teleconference pictures does not require close scrutiny of fine details like in broadcast or HDTV displays. Therefore, when using a test method originally designed for high quality pictures, some relaxation of test recommendations is permissible without affecting the results. Examples of such items are the number of evaluators, the spacing of tests by the same evaluator, the length of each test, and the exact values of picture luminance.

The setup used for the tests is shown on Figure 4.1. The two video tape recorders for the reference and processed tapes are each provided with a preview monitor needed by the recorder and switch operator. The 50 IRE pedestal produces a medium grey screen between the viewing sequences. The frame synchronizer minimizes switching transients. The viewing monitor for the evaluators is a high quality 19" color monitor.

Both tapes B and C consist of 16 sequences with durations ranging from 12 to 80 seconds, with most sequences lasting between 20 and 30 seconds. This results in a realistic crossection of the many types of scenes and presentation materials that may occur in a teleconference or videophone application. The large number of sequences is justified by the high importance of motion performance which must be evaluated in many different forms and contexts. Similar sequences verify the consistency of the ratings of each evaluator.

The tapes contain no audio, only a short "live" aural introduction was given at the start of the tests. Each sequence is first presented in its original form as reference, followed after a 3 second interval of medium grey background by the processed sequence. Immediately following is the 10 second scoring interval

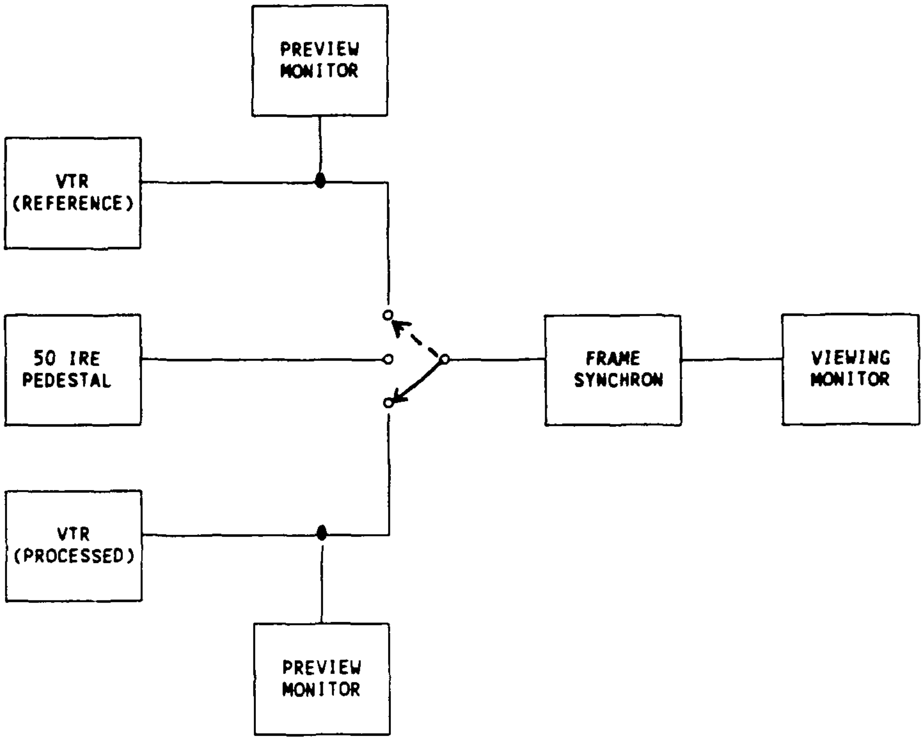


FIGURE 4.1
SUBJECTIVE TEST SETUP

which is visually identified with the sequence number to be scored. After a very short grey interval, the next reference sequence is presented. This timing arrangement including the tape recorder control functions is graphically shown on Figure 4.2. A slightly different arrangement is needed for the transition between parts B and C of the test tape which contains some unnecessary material. This arrangement is shown on Figure 4.3. The interval of about 30 seconds gives the evaluators a short rest and time to switch score sheets on their clip boards.

The tests were performed in a fully darkened room with light beige walls and controlled lighting. Four chairs for evaluators were provided in two rows, with the chairs in front about 4H and the chairs in the rear about 6H distance from the monitor screen where H is the displayed picture height. This arrangement allowed all observers an unobstructed view and complied with CCIR Rec 500-3. Light levels were kept close to the recommended values. A sketch of the room layout is shown on Figure 4.4. The numbers on each chair give the numbers of the evaluators occupying this position.

4.2 Test Personnel

Seven persons were available as evaluators complying with the stipulations

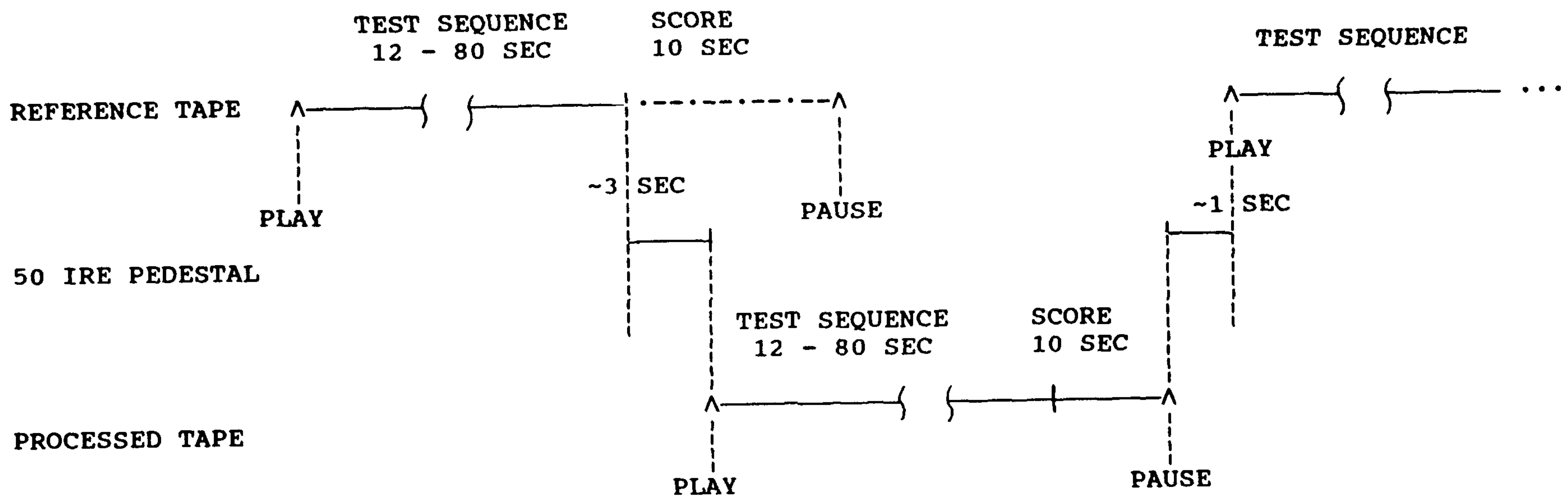


FIGURE 4.2
TEST TIMING

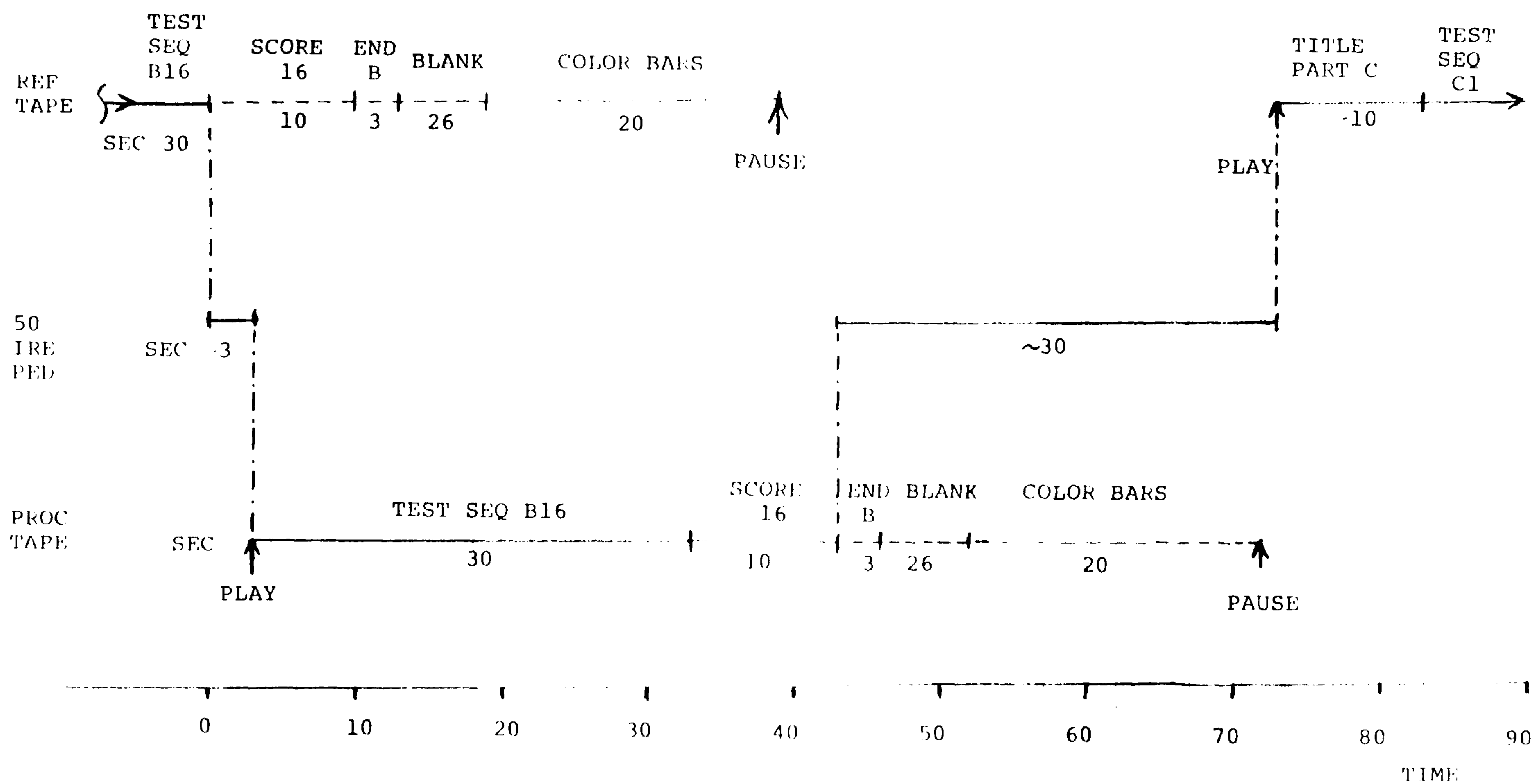


FIGURE 4.3

TIMING-TRANSITION BETWEEN PARTS B & C

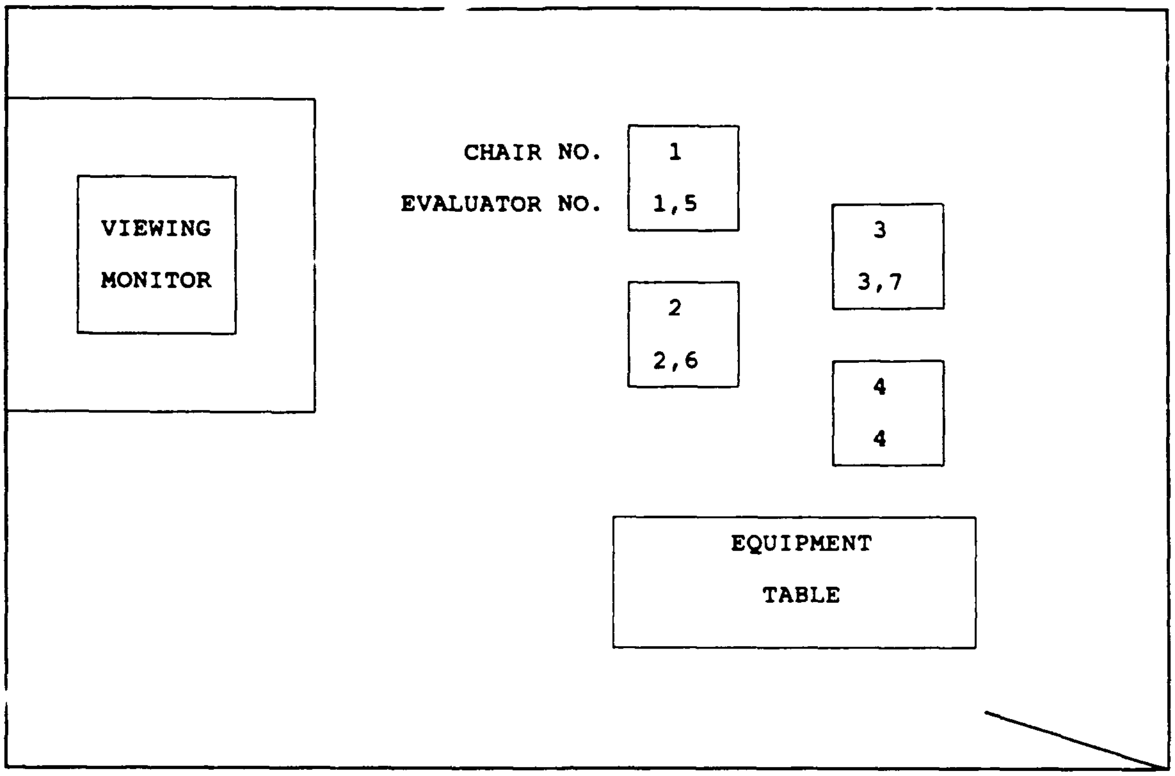


FIGURE 4.4
SUBJECTIVE TEST ROOM LAYOUT

of the Statement of Work. All had professional training, partly with technical background and some TV experience but none were experts in teleconferencing video codecs. Two groups, A and B, of four and three evaluators were formed, identified by numbers which also indicate their seating locations. Chair #4 was not occupied in group B. The two groups were necessary because it is not possible to place seven chairs within the range of acceptable viewing conditions.

Table 4-1 contains a list of evaluators and their qualifications, showing them to be representative of Government employees who would be likely users of a video phone or teleconferencing facility.

A filled-in sample of a score sheet used for each test is shown on Table 4-2 It identifies tape segment, evaluator and test by number and contains a guide for impairment grading and a line for each test sequence. Previous experience has shown that many evaluators prefer a somewhat flexible scale, therefore allowance is made for scoring between the official five grades. In some cases, an essentially continuous scale with 0.1 point divisions has been suggested but that much detail is unnecessary and may actually lead to confusion.

4.3 Scheduling

Proper scheduling of a test series like the one to be performed here is very

TABLE 4-1
TEST EVALUATORS

EVAL. NO.	GROUP	SEX	AGE	TITLE
1	A	F	33	Procurement Mgr.
2	A	M	50	Production Mgr.
3	A	F	32	Documentation Mgr.
4	A	F	58	Education Specialist
5	B	F	52	Financial Controller
6	B	F	43	Administration Mgr.
7	B	M	36	Project Engineer

important to ensure efficiency, fairness, and correct results. The task of performance grading is stressful, so enough rest between sessions is needed to minimize fatigue. Even so, the reactions of an evaluator often differ between beginning and end of a session. External influences may produce day-to-day variations. Consecutive tests by the same evaluator should be dissimilar to avoid any possible interaction.

The running times of the tapes including titles and scoring intervals are about 11 minutes for Tape B and 12 minutes for Tape C. Including the reference runs of the unprocessed tape and the additional short intervals, the total running time of both tapes reaches about 44 minutes, which is somewhat but not significantly beyond the suggested limits of CCIR REC 500-3. An interval of between 10 and 15 minutes between sessions had to be added for logistics purposes which brings the total time needed for each test up to about one hour.

Since the tests were alternated between two groups of evaluators and 7 tapes were available, a total of 14 tests had to be performed. The intervals between tests by the same group were shorter than suggested in CCIR REC 500-3 but experience has shown that this causes no adverse effect. The resulting schedule as implemented is shown on Table 4-3. This schedule was known only to the test director, none of the evaluators had any knowledge of codec operating modes or bit rates. The processed tapes were identified by numbers only. This schedule resulted in efficient use of facilities and personnel and satisfied all the above stated requirements. Tests No. 1 through 8 were performed on the first day, the remainder on the following day.

TABLE 4-2
EVALUATOR'S SCORE SHEET

CODEC EVALUATION FORM

EVALUATOR NO. 3 TEST NO. 13 DATE 10/31/91

IMPAIRMENT GRADING SCALE:

5: IMPERCEPTIBLE 4: PERCEPTIBLE BUT NOT ANNOYING
3: SLIGHTLY ANNOYING 2: ANNOYING 1: VERY ANNOYING

TAPE SEGMENT C

TEST SEQUENCE NO.	GRADE								
1	5		4		~~3~~		2		1
2	5		~~4~~		3		2		1
3	5		4		3		2		~~1~~
4	5		4		3		~~2~~		1
5	5		4		3		~~2~~		1
6	5		4		3		~~2~~		1
7	5		4		3		~~2~~		1
8	5		4		3		2		~~1~~
9	5		4		~~3~~		2		1
10	5		4		3		~~2~~		1
11	5		4		3		~~2~~		1
12	5		4		~~3~~		2		1
13	5		~~4~~		3		2		1
14	5		4		~~3~~		2		1
15	5		~~4~~		3		2		1
16	5		~~4~~		3		2		1

TABLE 4-3
TEST SCHEDULE

TEST NO.	EVALUATOR GROUP	TAPE NO.	CODEC FORMAT	BIT RATE
1	A	5	CIF	384
2	B	4	CIF	192
3	A	2	QCIF	192
4	B	6	CIF	768
5	A	7	CIF	1536
6	B	1	QCIF	128
7	A	3	CIF	128
8	B	7	CIF	1536
9	A	1	QCIF	128
10	B	3	CIF	128
11	A	6	CIF	768
12	B	2	QCIF	192
13	A	4	CIF	192
14	B	5	CIF	384

4.4 Results

The results of all subjective tests are listed on Tables 4-4 to 4-10, one table for each codec format and bit rate evaluated. The two bottom right numbers are the mean overall rating and its standard deviation. The individual scores were scrutinized to check them for overall consistency and to determine any erratic values. As expected, scores vary over a wide range, depending on test sequence contents and evaluator. Each evaluator obviously had to form his/her own interpretation of the impairment grades which is typical for all subjective tests. Some evaluators tend to give lower grades than others but none could be identified as being consistently the lowest or highest scorer. The various test sequences were deliberately designed to provide different levels of stress on the codec algorithm, therefore the variation of scores between sequences confirms that the

TABLE 4-4

QCIF128 CODEC EVALUATION SUMMARY

SEQ	EVALUATOR 1	2	3	4	5	6	7	MEAN	STD DEV
SEGMENT B									
1	1	4	1	1	2	2	1.5	1.79	.99
2	2	3	1	1	2	2	1.5	1.79	.65
3	2	2	1	2.5	1	3	1.5	1.86	.69
4	1.5	4	1	1.5	2	2	1	1.86	.95
5	1	3	2	1.5	1	1	1.5	1.57	.68
6	2.5	2	2	2	1	2	1.5	1.86	.44
7	2	3	1	1.5	1	1	1	1.50	.71
8	2.5	3	2	1	2	1	1	1.79	.75
9	3	2	3	1.5	1	2	2	2.07	.68
10	3	4	2	2	2	3	1.5	2.50	.80
11	3	2	3	2.5	1	3	1	2.21	.84
12	1	4	1	2	1	1	1.5	1.64	1.03
13	1	3	1	1.5	1	1	1	1.36	.69
14	3	3	3	2.5	2	3	2	2.64	.44
15	3	4	3	2.5	1	3	2	2.64	.87
16	2	2	1	1.5	1	2	1.5	1.57	.42
MEAN	2.09	3.00	1.75	1.75	1.38	2.00	1.44	1.92	.73
STD DEV	.77	.79	.83	.53	.48	.79	.35	.38	
SEGMENT C									
1	3	3	3	2.5	2	2	3	2.64	.44
2	2	1	2	3	2	3	2.5	2.21	.65
3	3	2	3	3	2	2	3.5	2.64	.58
4	1	2	1	2.5	1	1	2.5	1.57	.68
5	2	3	1	2.5	1	1	2	1.79	.75
6	2	4	1	3	2	2	2.5	2.36	.87
7	3	3	2	2	1	2	2	2.14	.64
8	1	2	1	2	2	2	2	1.71	.45
9	1	4	2	1	1	1	1	1.57	1.05
10	2	3	2	2	1	1	1.5	1.79	.65
11	2	2	1	2	1	1	1.5	1.50	.46
12	2	4	2	1.5	1	1	2	1.93	.94
13	3	4	1	2.5	2	1	2.5	2.29	.99
14	1	1	1	2	1	1	2	1.29	.45
15	3	4	3	2	2	3	2.5	2.79	.65
16	3	4	3	2	1	3	2.5	2.64	.87
MEAN	2.13	2.88	1.81	2.22	1.44	1.69	2.22	2.05	.70
STD DEV	.78	1.05	.81	.53	.50	.77	.59	.46	
TOTAL MEAN	2.11	2.94	1.78	1.99	1.41	1.85	1.83	1.99	.47

SUBJECTIVE TEST SCORES-128Kbps QCIF

TABLE 4-5

QCIF192 CODEC EVALUATION SUMMARY

SEQ	EVALUATOR 1	2	3	4	5	6	7	MEAN	STD DEV
SEGMENT B									
1	1	1	1	1	2	4	2	1.71	1.03
2	2	2	2	3	2	4	2	2.43	.73
3	2	2	1	1	1	3	2	1.71	.70
4	1	3	1	3.5	2	3	2	2.21	.92
5	1	1	1	3	1	2	1.5	1.50	.71
6	2	3	1	4	2	3	2	2.43	.90
7	1	1	2	1	1	1	2	1.29	.45
8	1	1	2	1	1	1	2	1.29	.45
9	1	3	2	4	2	1	3	2.29	1.03
10	1	1	2	2	2	2	2.5	1.79	.52
11	2	3	3	4	3	3	2	2.86	.64
12	1	3	2	1	2	1	1.5	1.64	.69
13	1	1	3	1	1	1	2	1.43	.73
14	2	2	1	4	2	3	3.5	2.50	.96
15	1	4	1	3	2	3	3	2.43	1.05
16	1	1	1	2	1	1	2.5	1.36	.58
MEAN	1.31	2.00	1.63	2.41	1.69	2.25	2.22	1.93	.76
STD DEV	.46	1.00	.70	1.24	.58	1.09	.53	.50	
SEGMENT C									
1	3	3	3	4	2	3	3.5	3.07	.56
2	2	3	2	2	2	3	3	2.43	.49
3	2	3	3	4	1	3	4	2.86	.99
4	1	2	2	2	2	2	2	1.86	.35
5	2	3	2	1	1	1	2	1.71	.70
6	2	3	3	2	2	2	3.5	2.50	.60
7	2	3	1	3	2	2	3	2.29	.70
8	1	3	2	1	2	3	2	2.00	.76
9	1	1	1	1	1	1	1	1.00	.00
10	2	2	3	1	1	3	2	2.00	.76
11	2	1	1	1	2	1	1.5	1.36	.44
12	2	2	3	3	1	2	2.5	2.21	.65
13	3	3	4	3	2	2	2.5	2.79	.65
14	1	1	2	3	1	2	2	1.71	.70
15	2	2	3	4	2	3	2.5	2.64	.69
16	3	1	4	3	1	3	2.5	2.50	1.04
MEAN	1.94	2.25	2.44	2.38	1.56	2.25	2.47	2.18	.63
STD DEV	.66	.83	.93	1.11	.50	.75	.76	.55	
TOTAL MEAN	1.63	2.13	2.04	2.40	1.63	2.25	2.35	2.06	.34

SUBJECTIVE TEST SCORES-192Kbps QCIF

TABLE 4-6

CIF128 CODEC EVALUATION SUMMARY

SEQ	1	2	3	4	5	6	7	MEAN	STD DEV
				EVALUATOR					
SEGMENT B									
1	1	1	1	1	2	3	1.5	1.50	.71
2	2	2	3	1	2	2	2	2.00	.53
3	1	2	2	1	2	4	2	2.00	.93
4	1	1	1	1	1.5	3	1	1.36	.69
5	1	2	2	1	1	3	1.5	1.64	.69
6	1	2	1	1	1	2	2	1.43	.49
7	1	1	1	1	1	1	2	1.14	.35
8	1	1	1	1	1	1	2.5	1.21	.52
9	1	3	2	1	1	2	3	1.86	.83
10	2	2	2	1	2	3	3	2.14	.64
11	2	2	2	2	2	3	3	2.29	.45
12	2	3	2	2	3	4	3.5	2.79	.75
13	1	2	1	1	1	2	2.5	1.50	.60
14	3	3	3	3	2	4	4	3.14	.64
15	2	3	2	2	1	3	3	2.29	.70
16	2	2	1	3	3	2	2.5	2.21	.65
MEAN	1.50	2.00	1.69	1.44	1.66	2.63	2.44	1.91	.64
STD DEV	.61	.71	.68	.70	.68	.93	.77	.54	
SEGMENT C									
1	3	2	3	4	3	4	4	3.29	.70
2	3	2	1	1	3	3	3.5	2.36	.95
3	3	3	3	4	3	4	4	3.43	.49
4	1	1	1	1	2	2	2.5	1.50	.60
5	1	2	1	1	1	3	2.5	1.64	.79
6	2	1	2	1	2	2	2	1.71	.45
7	2	3	2	3	2	2	2.5	2.36	.44
8	1	1	1	1	1	1	2.5	1.21	.52
9	1	1	2	1	1	2	1.5	1.36	.44
10	1	1	2	1	2	1	2	1.43	.49
11	2	2	2	1	1	1	2	1.57	.49
12	2	2	2	1	1	1	2.5	1.64	.58
13	3	2	3	2	2	2	3.5	2.50	.60
14	1	1	2	1	1	1	3	1.43	.73
15	3	3	4	4	4	3	4	3.57	.49
16	3	3	4	5	2	4	4	3.57	.90
MEAN	2.00	1.88	2.19	2.00	1.94	2.25	2.88	2.16	.61
STD DEV	.87	.78	.95	1.41	.90	1.09	.82	.84	
TOTAL MEAN	1.75	1.94	1.94	1.72	1.80	2.44	2.66	2.04	.37

SUBJECTIVE TEST SCORES-128Kbps CIF

TABLE 4-7

CIF192 CODEC EVALUATION SUMMARY

SEQ	EVALUATOR 1	2	3	4	5	6	7	MEAN	STD DEV
SEGMENT B									
1	3	1	4	3	2	4	4	3.00	1.07
2	2	1	2	3	2	4	3	2.43	.90
3	2	2	2	3.5	2	3	4	2.64	.79
4	2	3	1	3.5	2	4	2	2.50	.96
5	2	1	1	3	2.5	4	2	2.21	.99
6	3	2	3	3	3	3	4	3.00	.53
7	1	1	1	2	1	2	1	1.29	.45
8	2	3	2	2.5	2	3	2	2.36	.44
9	2	3	2	3	1.5	3	4	2.64	.79
10	2	2	1	3	3	4	3	2.57	.90
11	3	2	3	3.5	3	4	4	3.21	.65
12	3	4	3	3	3.5	3	3.5	3.29	.36
13	2	1	2	2	2.5	2	3	2.07	.56
14	3	2	4	3.5	3	4.55	4	3.44	.79
15	2	3	2	3	3	4.5	3	2.93	.78
16	2	3	3	3.5	2.5	3	3	2.86	.44
MEAN	2.25	2.13	2.25	3.00	2.41	3.44	3.09	2.65	.71
STD DEV	.56	.93	.97	.47	.64	.77	.91	.52	
SEGMENT C									
1	4	4	3	3.5	3.5	4	4.5	3.79	.45
2	3	2	2	2	2.5	4	3	2.64	.69
3	4	5	1	2.5	4.5	3.5	5	3.64	1.36
4	3	2	2	2	2	3	3	2.43	.49
5	2	1	2	2	2	2	4	2.14	.83
6	2	1	2	2	3	3	3	2.29	.70
7	3	3	2	2.5	2	4	4	2.93	.78
8	1	1	1	1.5	2	2	2.5	1.57	.56
9	1	4	3	2.5	2.5	3	3.5	2.79	.88
10	1	1	2	2	3	3	3.5	2.21	.92
11	2	2	2	2	2	4	4	2.57	.90
12	2	2	3	2	3.5	4	4.5	3.00	.96
13	3	4	4	2.5	3.5	4	5	3.71	.75
14	1	1	3	2	2.5	3	4	2.36	1.03
15	3	5	4	3	3.5	4	4.5	3.86	.69
16	2	5	4	3	3	4	5	3.71	1.03
MEAN	2.31	2.69	2.50	2.31	2.81	3.41	3.94	2.85	.81
STD DEV	.98	1.53	.94	.50	.73	.69	.77	.68	
TOTAL MEAN	2.28	2.41	2.38	2.66	2.61	3.43	3.52	2.75	.52

SUBJECTIVE TEST SCORES-192Kbps CIF

TABLE 4-8

CIF384 CODEC EVALUATION SUMMARY

SEQ	EVALUATOR 1	2	3	4	5	6	7	MEAN	STD DEV
SEGMENT B									
1	2.5	2	4	3	3	5	3.5	3.29	.92
2	3	2	2	4	4	5	4	3.43	1.05
3	4	4	2	2	3	4	3.5	3.21	.84
4	3	4	3	3	4	4	2.5	3.36	.58
5	2	3	1	3	4	4	3.5	2.93	1.02
6	4	4	2	2	4	4	3.5	3.36	.87
7	2	2	1	2	3	3	3	2.29	.70
8	3	1	2	3.5	4	4	3.5	3.00	1.04
9	3	4	4	5	3	4	4	3.86	.64
10	4	3	3	4	4	5	4	3.86	.64
11	4	4	3	3	4	5	3.5	3.79	.65
12	3	4	3	3.5	5	4	4.5	3.86	.69
13	2	3	2	2	4	4	4	3.00	.93
14	4	4	4	4	4.5	5	4.5	4.29	.36
15	3	5	4	3	3	4	4	3.71	.70
16	3	4	2	2.5	3	4	3.5	3.14	.69
MEAN	3.09	3.31	2.63	3.09	3.72	4.25	3.69	3.40	.77
STD DEV	.71	1.04	.99	.85	.61	.56	.50	.47	
SEGMENT C									
1	4.5	4	4	4.5	4	5	5	4.43	.42
2	4	4	3	2.5	3	4	4.5	3.57	.68
3	4	5	4	5	5	5	5	4.71	.45
4	3	3	2	3.5	3	4	3	3.07	.56
5	2	3	2	1.5	3	4	3	2.64	.79
6	2	3	2	2.5	4	4	3.5	3.00	.80
7	2	3	2	4	4	5	5	3.57	1.18
8	2	3	1	2	3	4	3	2.57	.90
9	3	2	2	4	2	3	2.5	2.64	.69
10	3	4	2	2.5	4	4	3.5	3.29	.75
11	2	3	1	1	3	3	3	2.29	.88
12	3	3	4	3	4	4	3	3.43	.49
13	4	3	4	2.5	5	5	5	4.07	.94
14	3	2	4	1	5	4	3.5	3.21	1.25
15	4	4	4	4	5	5	5	4.43	.49
16	4	3	4	5	4	5	5	4.29	.70
MEAN	3.09	3.25	2.81	3.03	3.81	4.25	3.91	3.45	.75
STD DEV	.87	.75	1.13	1.27	.88	.66	.94	.73	
TOTAL MEAN	3.09	3.28	2.72	3.06	3.77	4.25	3.80	3.42	.50

SUBJECTIVE TEST SCORES-384Kbps CIF

TABLE 4-9

CIF768 CODEC EVALUATION SUMMARY

SEQ	EVALUATOR 1	2	3	4	5	6	7	MEAN	STD DEV
SEGMENT B									
1	4	4	4	2.5	4.5	5	4	4.00	.71
2	5	4	4	4	4.5	5	4	4.36	.44
3	4	5	3	3.5	3	4	4.5	3.86	.69
4	4	4	4	4	4.5	4	2.5	3.86	.58
5	4	5	3	3.5	5	4	4.5	4.14	.69
6	5	5	4	4	4	4	4	4.29	.45
7	4	3	3	3.5	4	3	3	3.36	.44
8	4.5	3	4	4	4.5	4	4	4.00	.46
9	5		4	2	4.5	3	5	3.92	1.10
10	5	4	4	3	4.5	5	4.5	4.29	.65
11	5	5	4	4	5	5	4	4.57	.49
12	4	4	4	3.5	4.5	5	5	4.29	.52
13	4	4	4	4	4	4	4	4.00	.00
14	4	5	3	3.5	5	4	4.5	4.14	.69
15	4	5	4	3	3.5	4	4	3.93	.56
16	4	3	3	3.5	4.5	5	3.5	3.79	.70
MEAN	4.34	4.20	3.69	4.00	4.34	4.25	4.06	4.05	.57
STD DEV	.46	.75	.46	.57	.52	.66	.63	.28	
SEGMENT C									
1	5	5	4	3.5	4.5	5	5	4.57	.56
2	4.5	4	4	4	4	5	4.5	4.29	.36
3	5	4	4	3.5	4.5	4	5	4.29	.52
4	4	4	4	3.5	5	4	4.5	4.14	.44
5	4.5	3	3	3.5	3.5	3	4	3.50	.53
6	4	4	4	3.5	4.5	4	4	4.00	.27
7	5	4	4	3.5	4	5	4.5	4.29	.52
8	4	3	3	3	4	4	3.5	3.50	.46
9	4	4	3	3	3	4	4	3.57	.49
10	4	3	3	3	4.5	4	4	3.64	.58
11	4	4	3	3	3	3	3.5	3.36	.44
12	4.5	4	4	3.5	3	4	4.5	3.93	.49
13	5	5	4	3	4.5	4	4.5	4.29	.65
14	4.5	5	3	2.5	5	4	4.5	4.07	.90
15	4.5	4	5	3	5	4	5	4.36	.69
16	4	3	4	3	4.5	5	5	4.07	.78
MEAN	4.41	3.94	3.69	3.25	4.16	4.13	4.38	3.99	.54
STD DEV	.40	.66	.58	.35	.68	.60	.48	.36	
TOTAL MEAN	4.38	4.07	3.69	3.63	4.25	4.19	4.22	4.06	.30

SUBJECTIVE TEST SCORES-768Kbps CIF

TABLE 4-10

CIF1536 CODEC EVALUATION SUMMARY

SEQ	EVALUATOR 1	2	3	4	5	6	7	MEAN	STD DEV
SEGMENT B									
1	4.5	4	4	3	5	5	4	4.21	.65
2	5	5	5	5	5	5	4.5	4.93	.17
3	5	4	4	5	4	5	4.5	4.50	.46
4	4	5	5	4	4	5	3.5	4.36	.58
5	4	4	4	4	4.5	5	5	4.36	.44
6	4	4	4	4	3	4	4.5	3.93	.42
7	5	3	4	2	4	5	4	3.86	.99
8	4.5	4	5	3	3	5	4	4.07	.78
9	4.5	5	4	4	3	5	5	4.36	.69
10	5	5	4	3	4	5	4.5	4.36	.69
11	5	5	4	4	5	5	5	4.71	.45
12	4	4	4	2	4.5	4	5	3.93	.86
13	4	4	4	3	5	5	4.5	4.21	.65
14	4.5	4	5	4	5	5	5	4.64	.44
15	5	5	5	4	4	4	4.5	4.50	.46
16	4	4	4	5	3	4	5	4.14	.64
MEAN	4.50	4.31	4.31	3.69	4.13	4.75	4.53	4.32	.59
STD DEV	.43	.58	.46	.92	.76	.43	.45	.29	
SEGMENT C									
1	5	4	4	4	5	5	5	4.57	.49
2	4.5	4	4	4	5	5	4.5	4.43	.42
3	4.5	5	5	5	5	5	5	4.93	.17
4	4	4	4	4	5	5	4.5	4.36	.44
5	4	3	4	3	4	4	4	3.71	.45
6	4.5	4	4	3	4	4	5	4.07	.56
7	5	4	4	5	5	5	4.5	4.64	.44
8	4.5	3	4	3	5	4	5	4.07	.78
9	4	2	4	2	4	4	3.5	3.36	.87
10	5	4	4	3	4	5	4	4.14	.64
11	4	4	5	2	4	4	4.5	3.93	.86
12	4	4	4	3	3	4	4.5	3.79	.52
13	4.5	4	5	3	4	5	5	4.36	.69
14	4.5	4	5	3	4	5	4.5	4.29	.65
15	4	5	5	4	4	5	5	4.57	.49
16	5	5	5	5	4	4	5	4.71	.45
MEAN	4.44	3.94	4.38	3.50	4.31	4.56	4.59	4.25	.56
STD DEV	.39	.75	.48	.94	.58	.50	.44	.40	
TOTAL MEAN	4.47	4.13	4.35	3.60	4.22	4.66	4.56	4.28	.34

SUBJECTIVE TEST SCORES-1536Kbps CIF

test tape is serving its intended purpose.

The results were computed for tape segments B and C separately and also totaled for each complete test. This is desirable because rendition of details and distortions caused by motion, panning or zooming tend to have different effects on the judgement of the various evaluators. Only the total scores were used for overall results and correlation.

The individual test scores show the variations which have to be expected with all subjective tests but there are no obvious discrepancies. The scores show no influence of sex or age of the evaluators or of the test schedule. All standard deviation values are low, with only a few slightly higher than 1. The total scores increase consistently with bit rate as shown on Figure 4.5. At the two common bit rates, CIF scores are slightly higher than QCIF. This indicates that the low definition of QCIF is even more disturbing than the deteriorated motion performance of CIF at these low bit rates. There is generally not much difference between the scores for graphics with motion and persons with limited motion, but there is a tendency for the scores for persons to be slightly higher at low bit rates which reverses as the rates are increased. This indicates that severe jerkiness may be more annoying in motion connected with graphics than with persons.

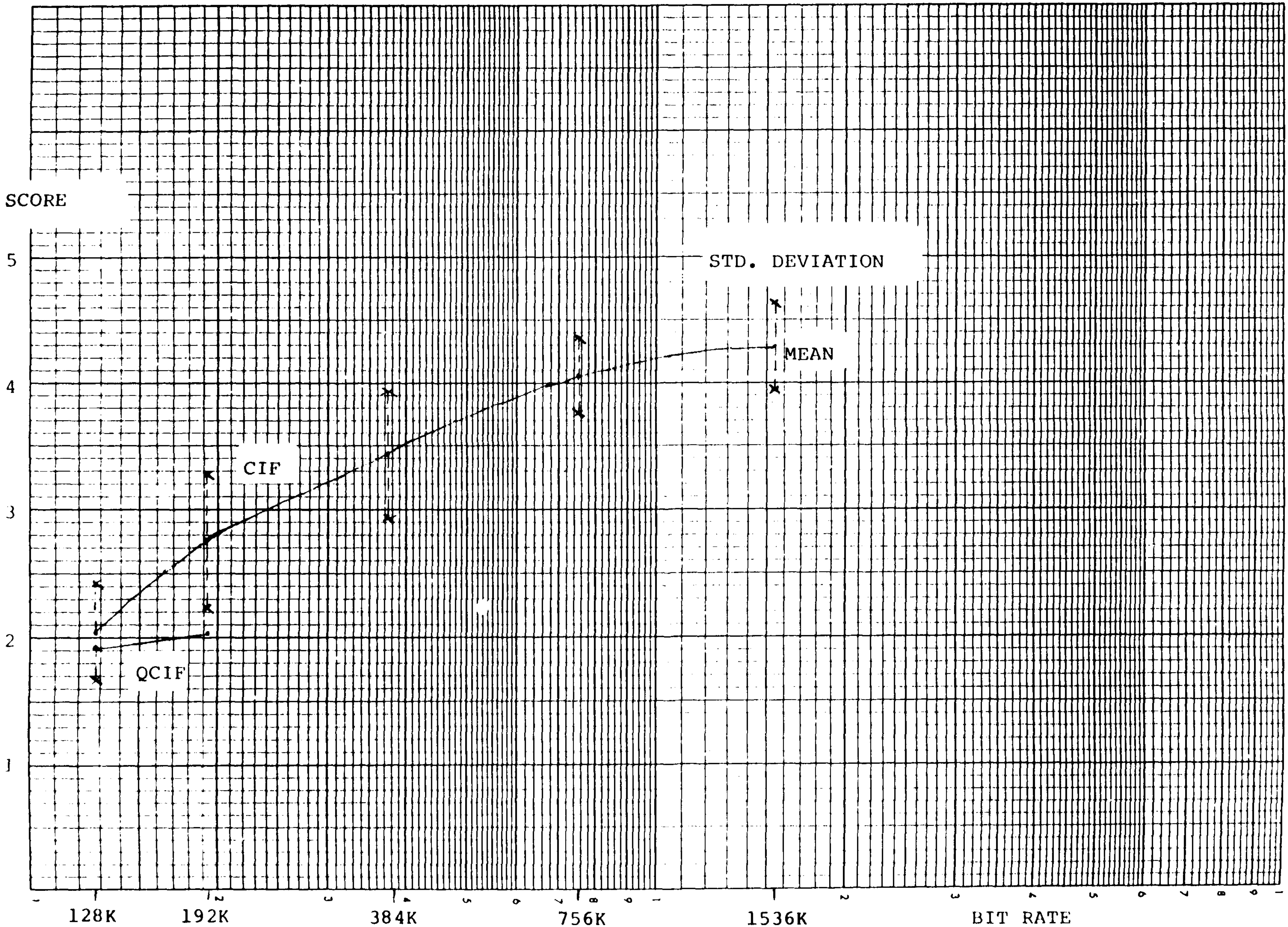


FIGURE 4.5 SUBJECTIVE TEST SCORES

5.0 OBJECTIVE TESTS

5.1 Test Methodology

The objective test methodology which was previously developed and used successfully as described in NCS-TIB-90-6 and NCS-TIB-91-2 was implemented essentially unchanged. Most tests were performed with the rotating wheel test pattern. Each series of tests used patterns with two different spoke widths which confirmed the logical relations between test pattern parameters and results.

Though most codec designs feature the same elements and functions, there are enough possible differences in their implementation that some adaptation of the test methods is necessary to take all variations in codec performance into account. Some of these variations from previously tested codecs can be explained by the specific design features stipulated by H.261 but others will require further investigation. All these factors are likely to have an impact on the correlation of subjective and objective test results.

5.2 Temporal Response

The test setup used for this measurement was described previously in NCS-TIB-91-2, but for completeness is included here as Figure 5.1. The processed tape

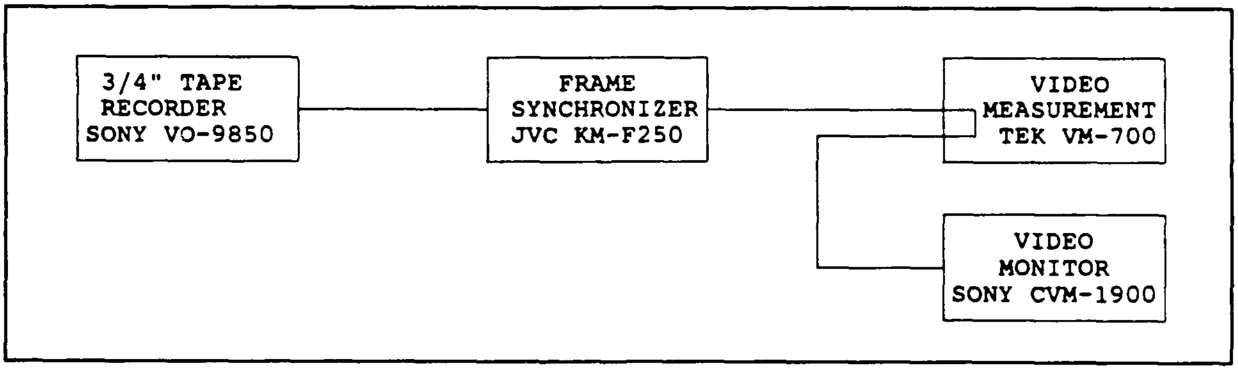


FIGURE 5.1
TEST SETUP
TEMPORAL FREQUENCY RESPONSE

is played on the VO-9850 Tape Recorder which allows manual tape advance and identification of each frame by means of a frame counter. The KM-F250 Frame Synchronizer re-generates the degraded synch waveform of the still picture from the tape recorder and makes it compatible with the VM-700 Video Measurement Set. This equipment allows amplitude measurement at any exactly defined point in the picture with an accuracy of better than one IRE unit. It also makes it possible to recognize all temporal advances of the test pattern. An amplitude measurement

is made after every advance even if there has been no change in amplitude. Measurements are taken over about 4 to 10 temporal cycles, depending on pattern detail, rotation speed and the amount of recognized amplitude variations.

Though all numerical measurements are made on the VM-700 Video Measurement Set, the video monitor performs more functions than merely identification of the test pattern. Any rotating spoke pattern can produce temporal aliasing on a TV monitor, indicated by an apparent rotation reversal. Since the codec under test extensively uses frame repetition, aliasing occurs at relatively low rotation speeds. The aliasing limit can also be recognized as the point where the number of frames per spoke (see Table 3-3) becomes equal to the frame repetition rate. This point represents the upper limit of valid measurements, at higher temporal frequencies the results become meaningless.

Another feature not previously observed is also connected with frame repetition. At relatively high repetition rates, the advance between two consecutive positions of the spoke pattern is not made in a single "jump" but there are one or two intermediate conditions when both positions are visible with reduced amplitude. When combined with temporal aliasing, an almost uniformly grey picture may occasionally result. These intermediate conditions are also easily recognized on the VM-700 displays as shown on Figures 5.2 to 5.5. The amplitudes measured at these points are not connected with an actual frame advance in the transmitted picture. This condition has not been observed previously and its cause is presently unknown. It may be connected with the conversion between interlaced and sequential scan which is unique in the H.261 codec.

Initial visual inspection showed that due to the wide range of bit rates not all test patterns produced significant and usable results. Patterns with wide spokes and low speeds do not provide sufficient challenge at high bit rates while narrow spokes and high speeds generally far exceed the codec capability at low bit rates. Therefore, specific groups of patterns were selected for numerical evaluation of each bit rate and codec format. A sample of such a selection is shown on Table 5-1 which is simply a marked-up copy of Table 3-3.

Using the test setup of Figure 5.1, an amplitude measurement is made following each test pattern advance. The results are recorded on a data sheet, a filled-in sample of which is shown on Table 5-2. The frame numbers marked with an asterisk do not represent an actual frame advance and are used only in the computation of RMS amplitude. The frame numbers and amplitudes are processed in a computer, resulting in a printout, a sample of which is shown on Figure 5.6. The RMS amplitude number on this figure represents one value of temporal response. Analyzing various rotation speeds of the same test pattern under identical test conditions results in a series of points on a temporal response curve. As indicated on Table 5-1, portions of two different spoke width patterns were analyzed for every bit rate. Thus, as shown on Figure 5.7, two temporal response curves can be drawn for every bit rate and operating mode. This demonstrates the dependence of the results on the test pattern. It does not seem necessary to draw

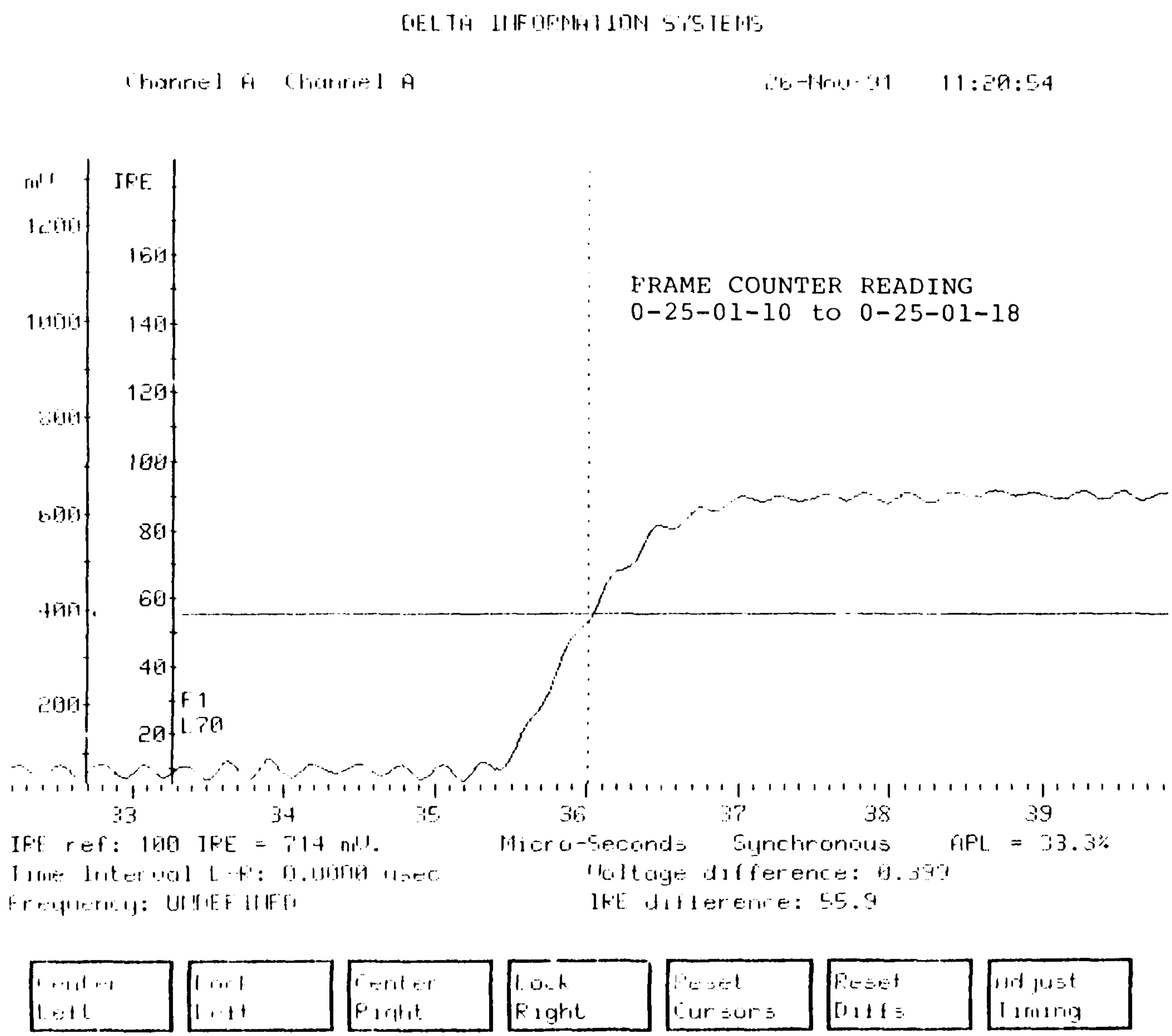


FIGURE 5.2 VM-700 DISPLAY-BEFORE FRAME ADVANCE

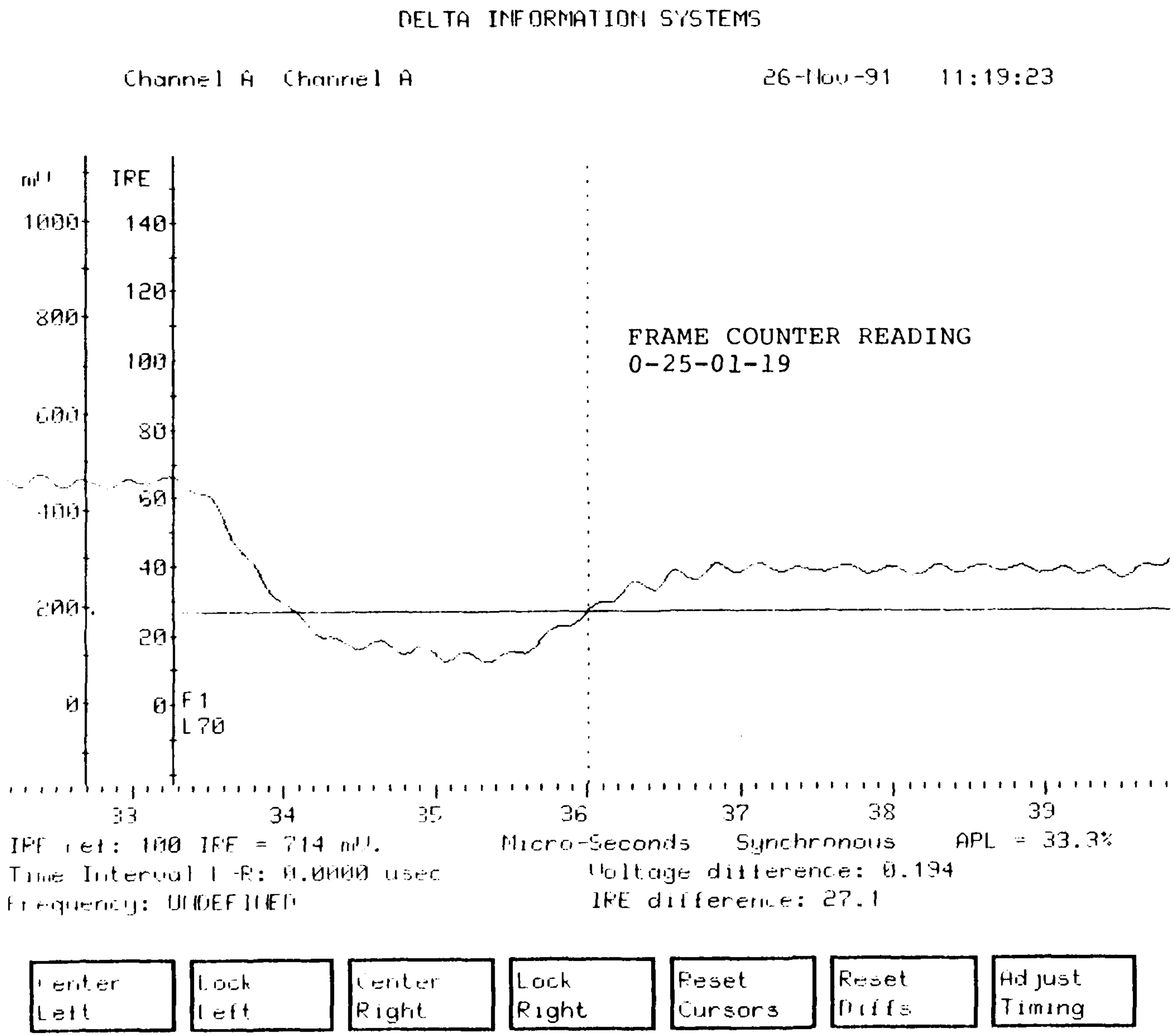


FIGURE 5.3 VM-700 DISPLAY-INTERMEDIATE CONDITION

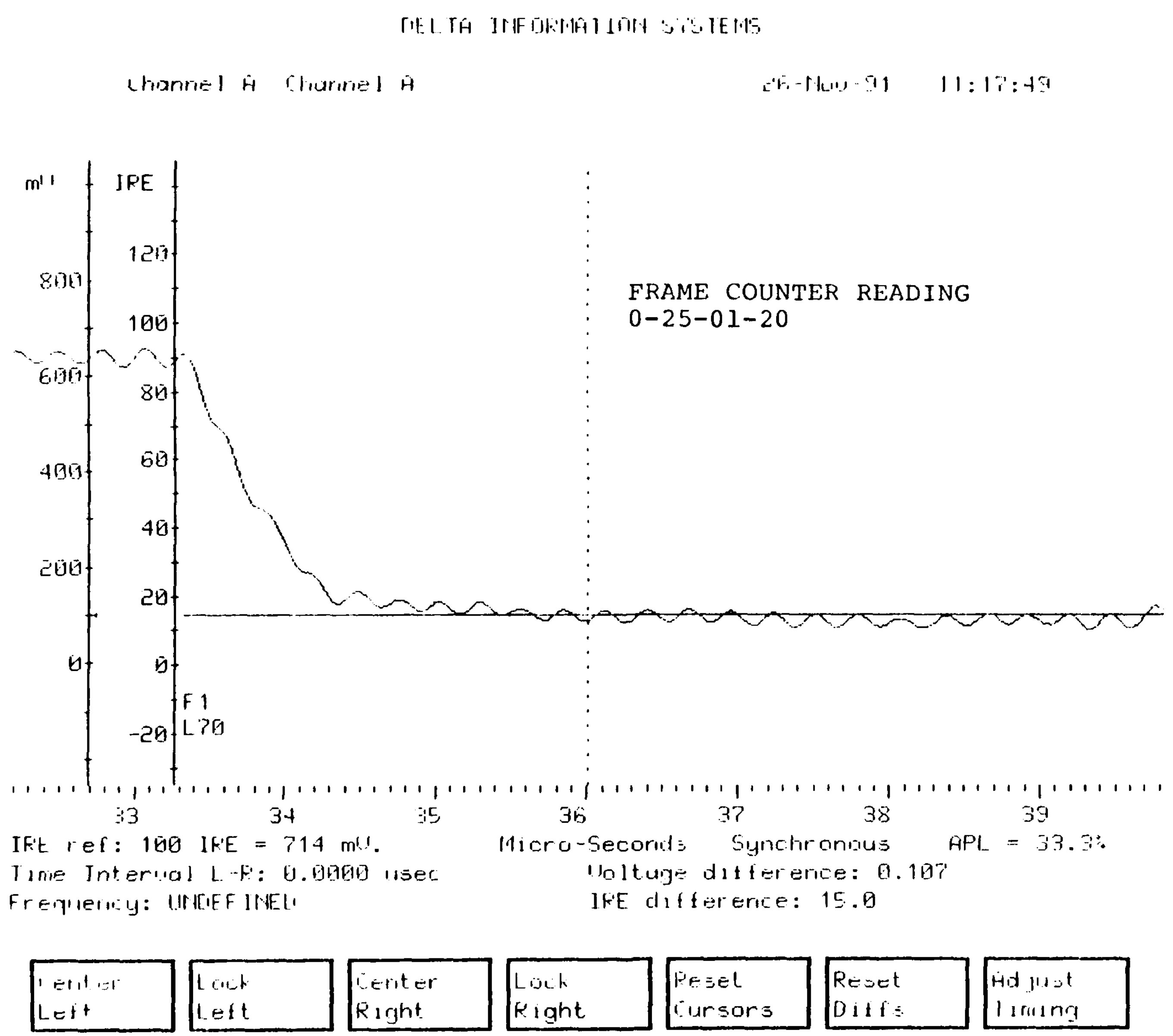


FIGURE 5.4 VM-700 DISPLAY-AFTER FRAME ADVANCE

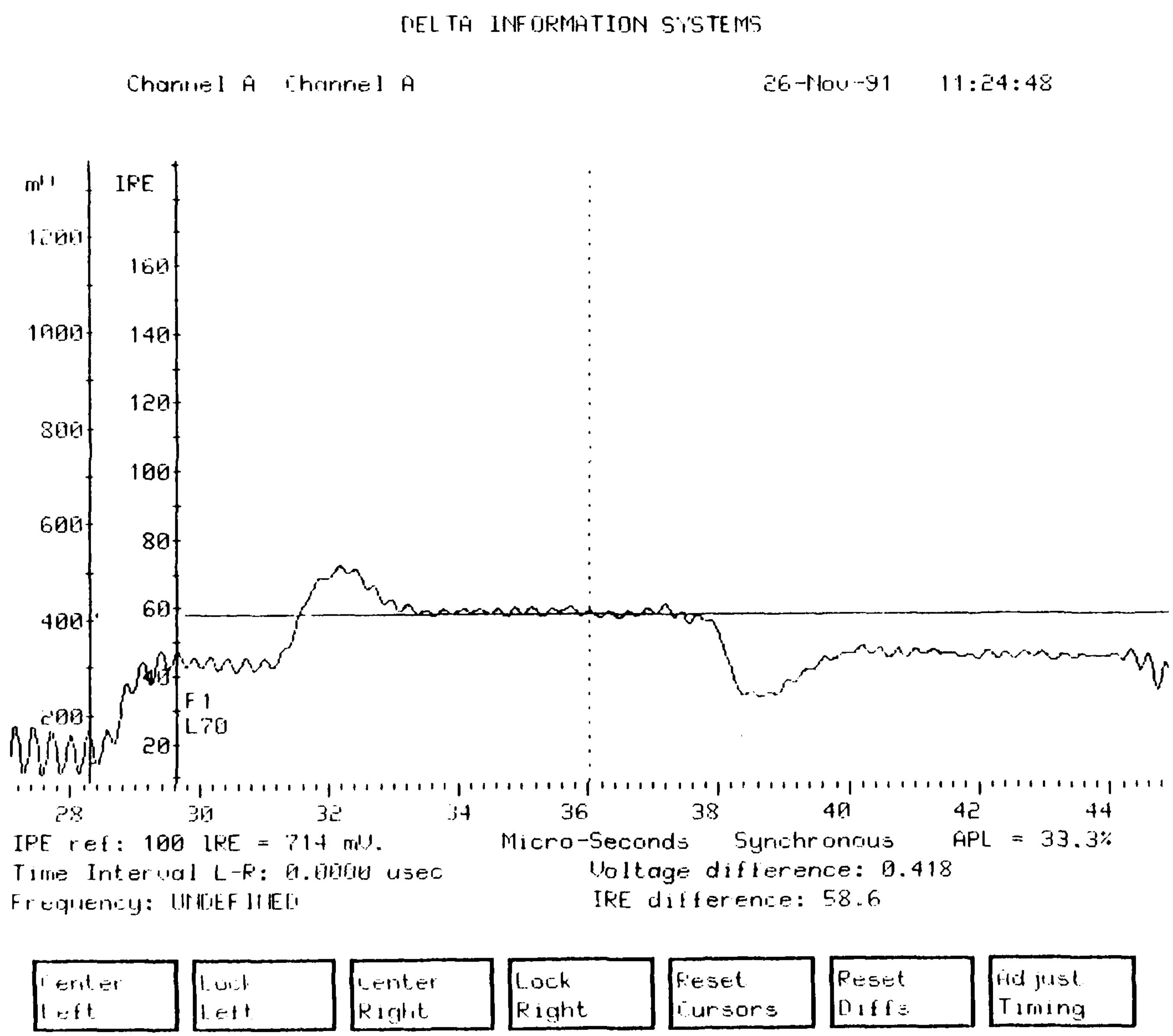


FIGURE 5.5 VM-700 DISPLAY-INTERMEDIATE CONDITION WITH ALIASING

TABLE 5-1
ROTATING WHEEL PATTERNS
SELECTED FOR QCIF 128

TEST PATTERN NO.	SPOKE WIDTH (DEGREES)	ROTATION SPEED FRAMES/ REVOLUTION	DEGREES/ SECOND	TEMPORAL FREQUENCY (CPS)	FRAMES/ SPOKE	% PIXEL CHANGE/ FRAME	% BLOCK CHANGE/ FRAME
1	30	540	20	0.33	45	2.2	18
2	30	360	30	0.50	30	3.3	
3	30	240	45	0.75	20	5.0	
4	30	180	60	1.00	15	6.7	22
5	30	144	75	1.25	12	8.3	
6	30	120	90	1.50	10	10.0	
7	30	90	120	2.00	7.5	13.3	34
8	30	72	150	2.50	6	16.7	
9	30	60	180	3.00	5	20.0	43
10	18	720	15	0.42	36	2.8	31
11	18	540	20	0.55	27	3.7	
12	18	360	30	0.85	18	5.6	
13	18	240	45	1.25	12	8.3	
14	18	180	60	1.67	9	11.1	42
15	18	144	75	2.10	7.2	13.9	
16	18	120	90	2.50	6	16.7	
17	18	90	120	3.33	4.5	22.2	57
18	10	720	15	0.75	20	5.0	50
19	10	540	20	1.00	15	6.7	54
20	10	360	30	1.50	10	10	
21	10	240	45	2.25	6.7	15	
22	10	180	60	3.00	5	20	70
23	10	144	75	3.75	4	25	75

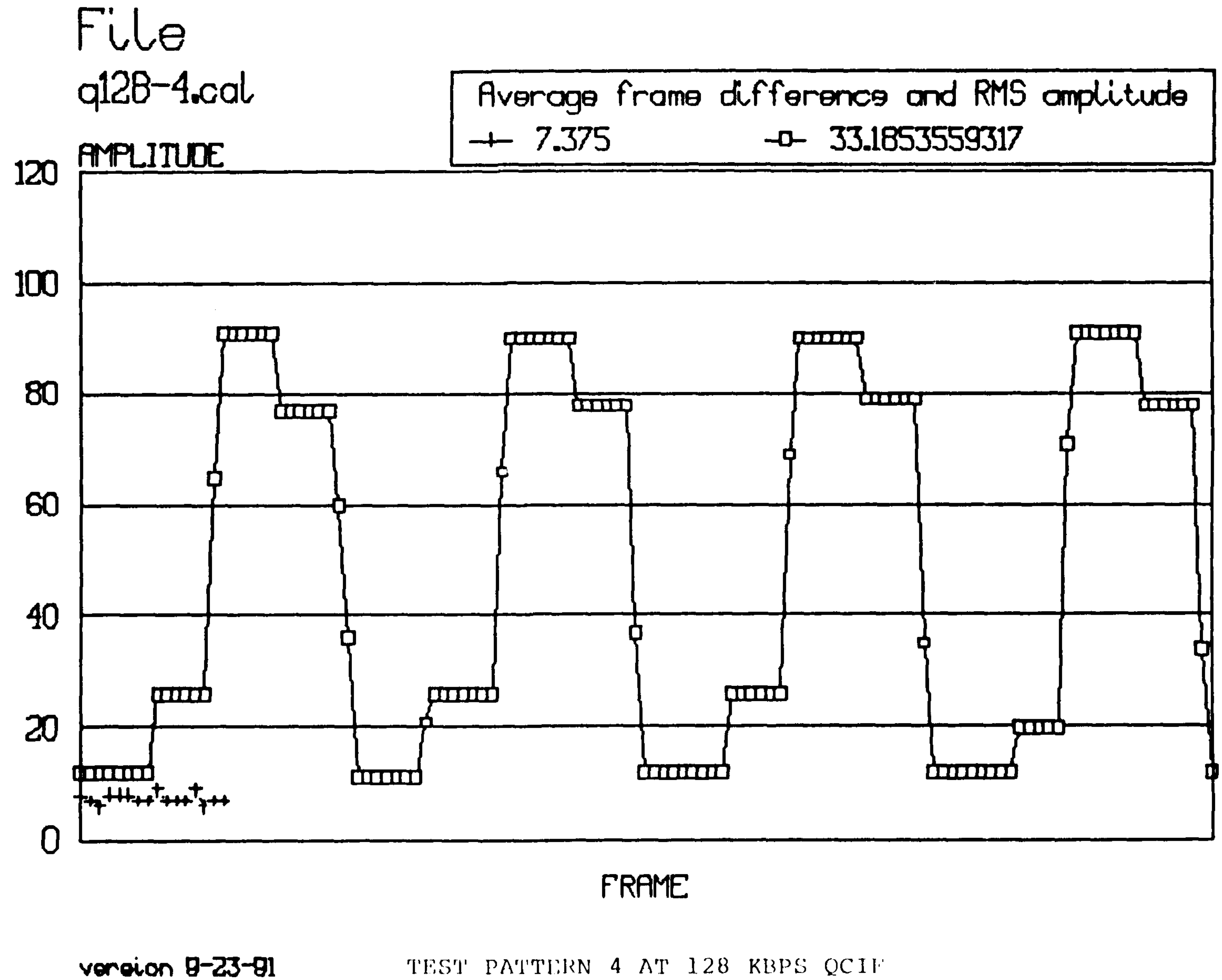
TABLE 5-2
AMPLITUDE MEASUREMENT DATA SHEET

CODEC: QCIF RATE: 128 PATTERN: 3 LINE: 70 POS: 36/MICROSEC

FR	AMP	FR	AMP	FR	AMP	FR	AMP	FR	AMP	FR	AMP
1	10	84	14								
8	14	91	14								
16	23	98	25								
23*	33	105*	44								
24*	68	106*	76								
24	90	106	90								
30	90	112	90								
36	73	118	50								
42	12	124*	36								
50	15	125	11								
58	26	132	16								
65*	45	140	55								
65*	73	145	90								
65	91	152	90								
71	91	158*	55								
77	76	158	50								
83*	60	166	12								
84*	29										

all such possible curves because they all are very similar. The average results are summarized on Table 5-3.

The results show that in contrast with previously examined proprietary designs, this P x 64 codec has a temporal response which is relatively independent of bit rate and temporal frequency. This seems primarily due to the fact that the basic parameters of the encoding algorithm are fixed. The main variable parameter is the quantization accuracy of the video signal which influences the signal-to-noise ratio but does not have primary impact on the ability of the codec to faithfully reproduce a moving black-white transition. Motion compensation contributes to this performance feature, the technique used in this codec may be particularly



TEST PATTERN 4 AT 128 KBPS QCIF

FIGURE 5.6 SAMPLE PRINTOUT

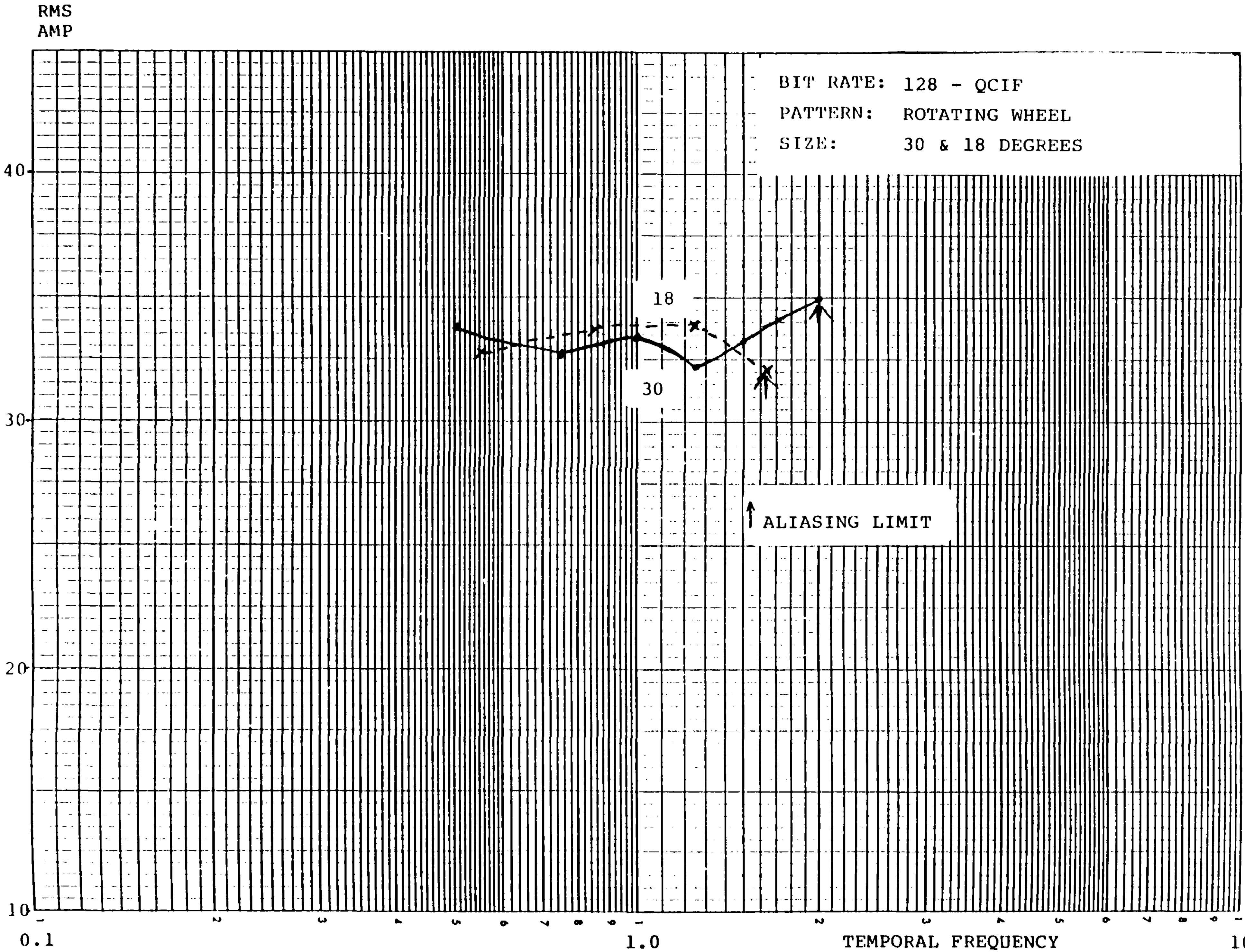


FIGURE 5.7 TEMPORAL FREQUENCY RESPONSE

TABLE 5-3
AVERAGE TEMPORAL RESPONSE

MODE & BIT RATE		Q128	Q192	C128	C192	C384	C768	C1536
SPOKE WIDTH	30	33	33	33	34	--	--	--
	18	32	31	32.5	33	32	32	32.5
	10	--	--	--	--	27	28	27.5

effective with the rotational motion of the test patterns. The width of the rotating spokes has a noticeable and consistent effect on the temporal response.

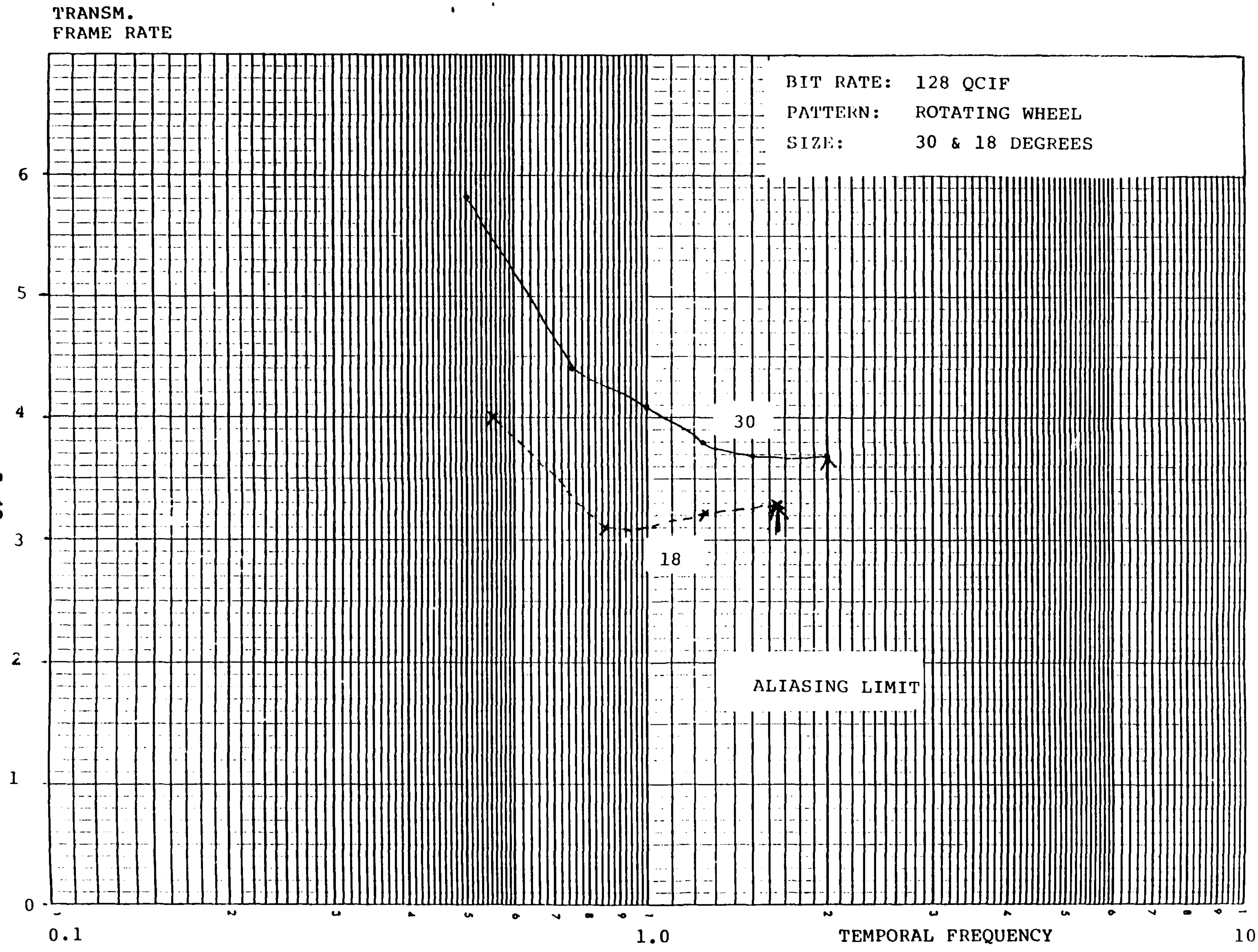
5.3 Transmitted Frame Rate

The H.261 Standard calls for means to restrict the maximum picture rate by having several non-transmitted picture between transmitted ones. This feature is being widely used and is displayed as part of the print-out on Figure 5.6 as the average difference between frame advances over the whole range of measurements. Though each frame repetition must be a whole number, averaging over several frames generally produces a fractional value. The transmitted frame rate is obtained by dividing 30 by the frame repetition rate.

Figure 5.8 shows two sample curves of transmitted frame rate vs temporal frequency. Table 5-4 gives the range of the results for all bit rates and operating modes. For the 18° spoke width, it also gives the average of all measured values. It makes it obvious that the transmitted frame rate varies over a wide range and is closely related to bit rate, operating mode and temporal frequency. The codec under test uses it as the main means to limit the transmission bit rate.

TABLE 5-4
TRANSMITTED FRAME RATES

MODE & BIT RATE			Q128	Q192	C128	C192	C384	C768	C1536
SPOKE WIDTH	30		5.8-3.7	6-5	3.9-1.8	3.9-1.8	---	---	---
	18	RANGE	4 - 3.1	4.7-3.8	2.7-1.5	2.3-1.5	9.7-5.4	30-13.4	30
		AVG.	3.4	4.2	2	2	6.7	23.3	30
	10		---	---	---	---	8.4-5.2	15.5-13.6	30-27.5


TRANSM.
FRAME RATE
BIT RATE: 128 QCIF
PATTERN: ROTATING WHEEL
SIZE: 30 & 18 DEGREES
30
18
ALIASING LIMIT
6
5
4
3
2
1
0
0.1
1.0
10
TEMPORAL FREQUENCY

5.4 Image Update Time

In videoconferencing and videophone systems there is typically an "update" mode of operation when the output image must change to a totally new picture as rapidly as possible. For example, this mode occurs when a scene cut is introduced. In this case, an artifact appears which is particularly visible at low bit rates. The artifact may take the form of blocking, blurring, retention of the previous picture, etc.

Measurements of image update time were performed in both CIF and QCIF modes using the scene cut patterns A-1, A-13, and A-25 as shown on Table 3-4 and the test setup of Figure 5.1. Switching from an all black screen to the white circle pattern with a switching interval of 4 seconds (120 frames) gave the best results. Update of the image occurs usually in up to 5 discrete steps, but is uniform over the whole picture.

The measurement results are summarized in the four simple graphs on Figure 5.9. The update time is defined as the number of frames necessary after switching to achieve 100% output level. This definition gives the widest range of useful data which are fully consistent in their relation to other system parameters. The update time always increases at lower data rates. Increased challenge on the codec algorithm produced by smaller circles has a noticeable effect at CIF. At QCIF, the update time is lower but essentially independent of circle size, therefore only one average line is shown. At higher data rate CIF, the update time is trivially low.

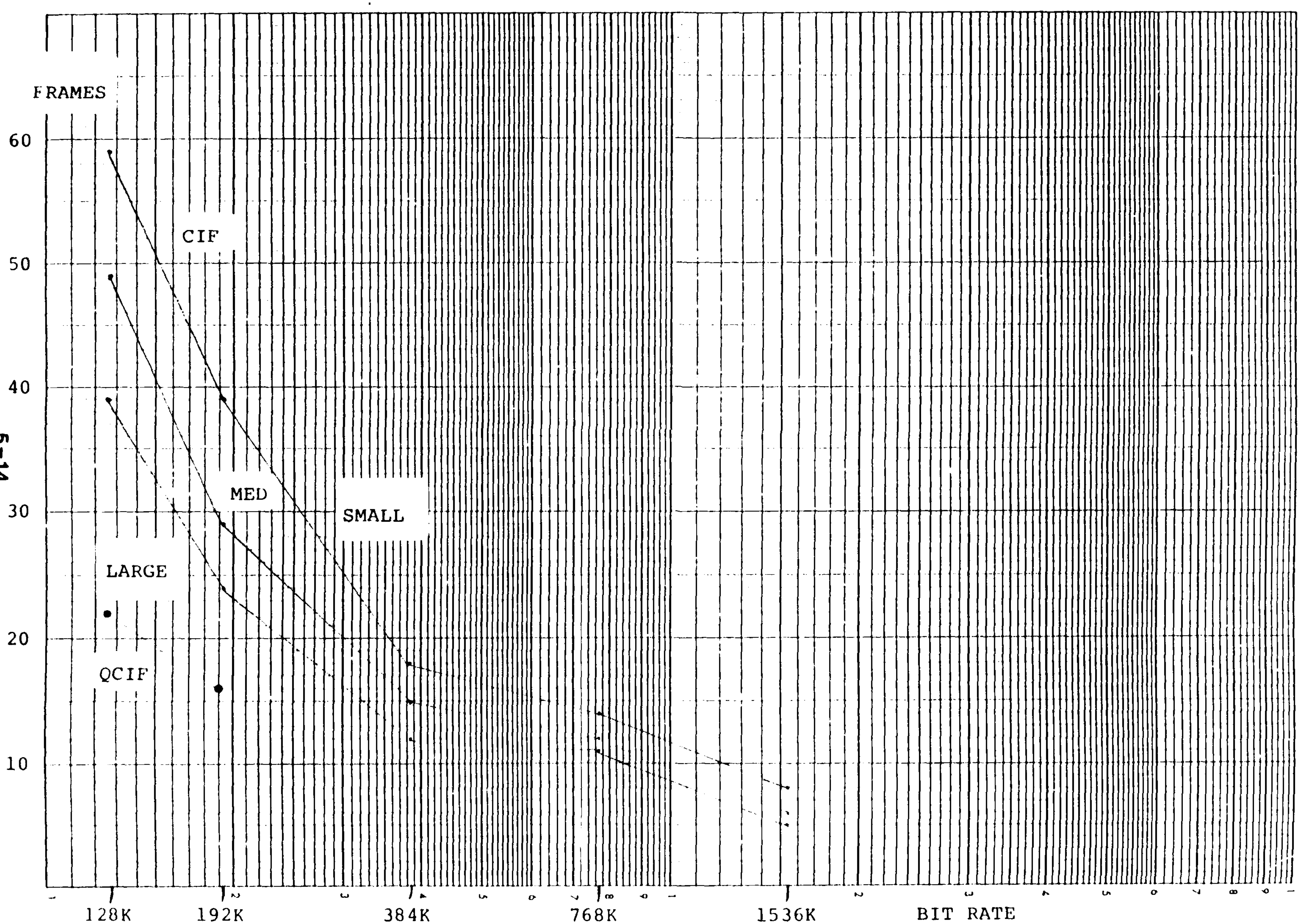


FIGURE 5.9 IMAGE UPDATE TIMES

6.0 CORRELATION

The first effort to establish correlation of subjective and objective test results was described in NCS-TIB-91-6. The same test methodology was used in this program and the same basic factors are applicable. In spite of some refinements, it is still not possible to objectively describe codec performance with one simple number, and at present it cannot be predicted when this will be achieved. Indeed, comparison of the results of Section 5 with those of previous programs obtained on different codecs shows that codec algorithms and other design details determine which parameters are critical for objective assessment of codec performance. Therefore, some rather arbitrary choices are still necessary and codec performance can be described objectively only with numerical values of several independent parameters.

Since the 18° spoke pattern was used for all evaluated bit rates at CIF and QCIF it was also used for all correlation evaluations of temporal response and transmitted frame rate. The limits imposed by temporal aliasing made it impossible to select values at a single temporal frequency. Since data were available at QCIF for only two bit rates, correlation evaluation was limited to CIF. The data on Tables 4-4 to 4-10, 5-3, 5-4 and Figure 5.9 were used and produced the results shown on Figures 6.1 to 6.3.

Figure 6.1 shows that temporal response is essentially unchanged over the full range of subjective scores. This indicates that the wide variations of the evaluations are based on other performance parameters. Evidently, this codec design does not allow conventional motion artifacts such as picture blurriness to limit the transmitted bit rate to a low value. A very efficient motion compensation design may be the main means to accomplish this result.

Figure 6.2 displays a wide range of transmitted frame rate which correlates well with the subjective scores. The number used for transmitted frame rate is the average of the values for all temporal frequencies used for measurement with the specific pattern and bit rate as shown on Table 5-4. It is evident that this codec design depends mainly or entirely on non-transmission of frames to reduce the transmitted bit rate. It appears that at the lowest bit rate, the number of non-transmitted frames reaches a pre-set maximum.

Figure 6.3 shows excellent correlation of the subjective scores with image update time. The results with the most and least challenging patterns (small and large circles) produce almost exactly parallel curves. Though in actual teleconference scenes, image update time manifests itself directly only when switching between very dissimilar pictures, it can also serve as a general measure of motion performance. This fact increases the significance of this parameter and its correlation with subjective scores.

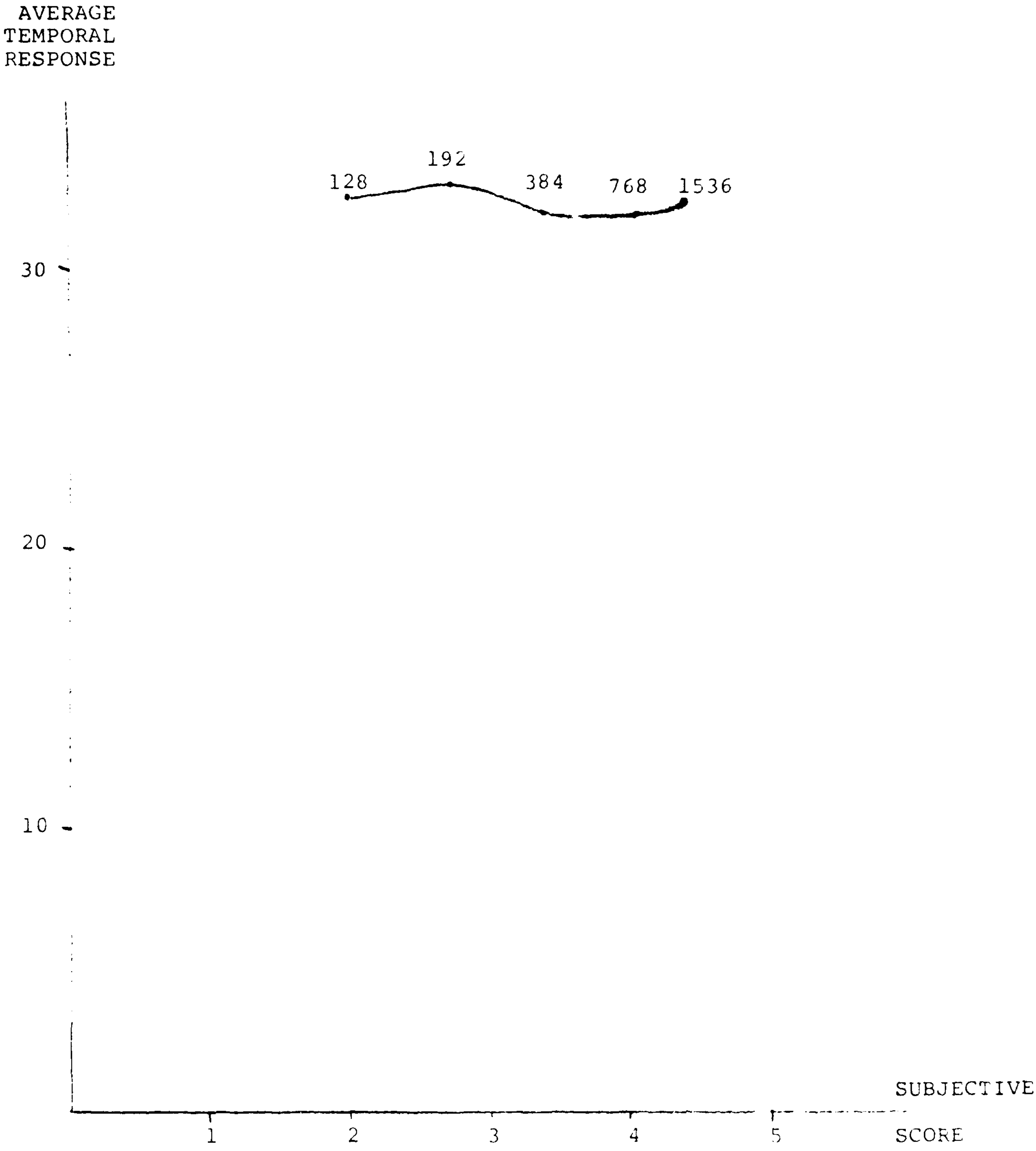


FIGURE 6.1 CORRELATION-SUBJECTIVE SCORE
vs. TEMPORAL RESPONSE

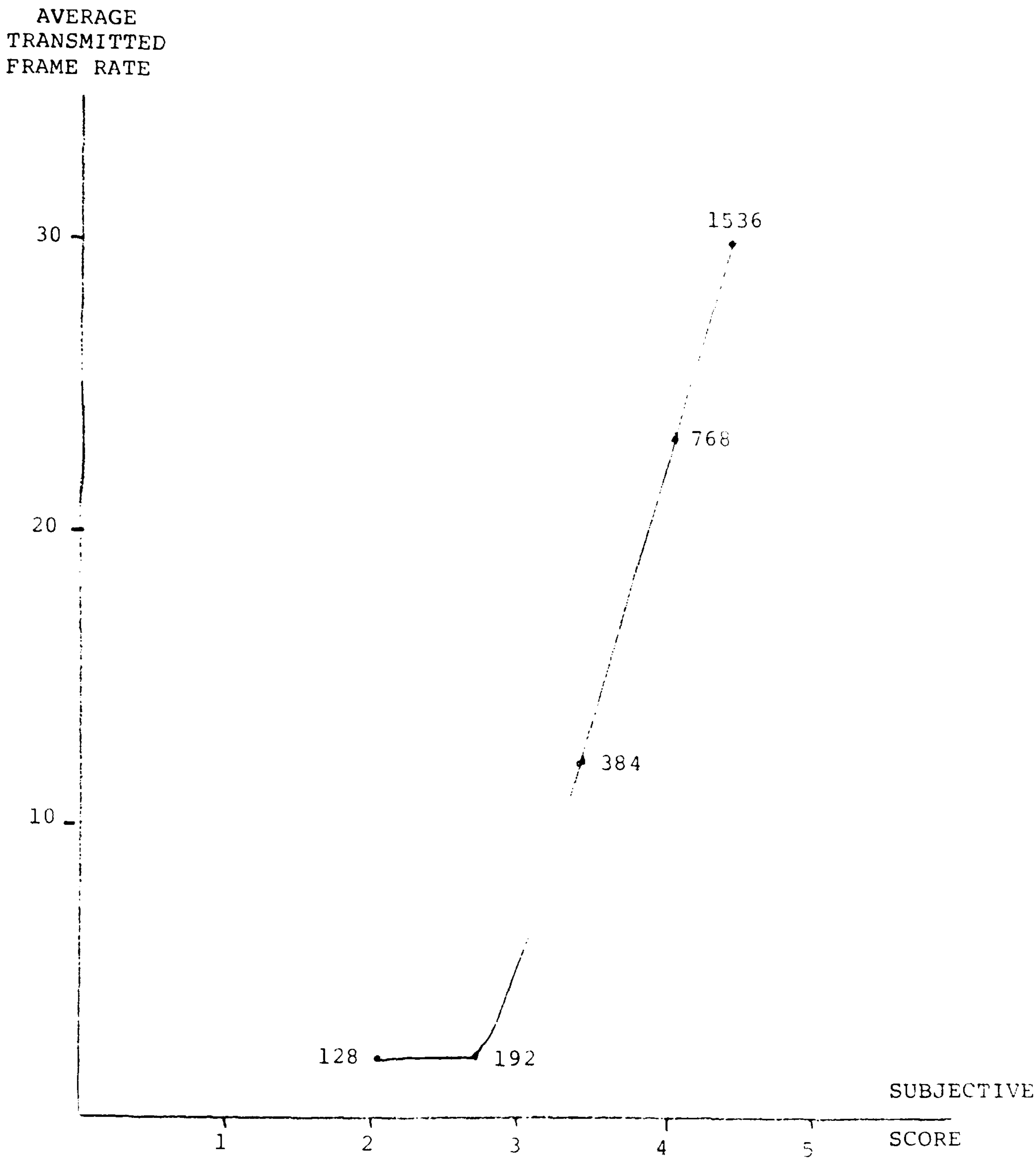


FIGURE 6.2 CORRELATION-SUBJECTIVE SCORE
vs. TRANSMITTED FRAME RATE

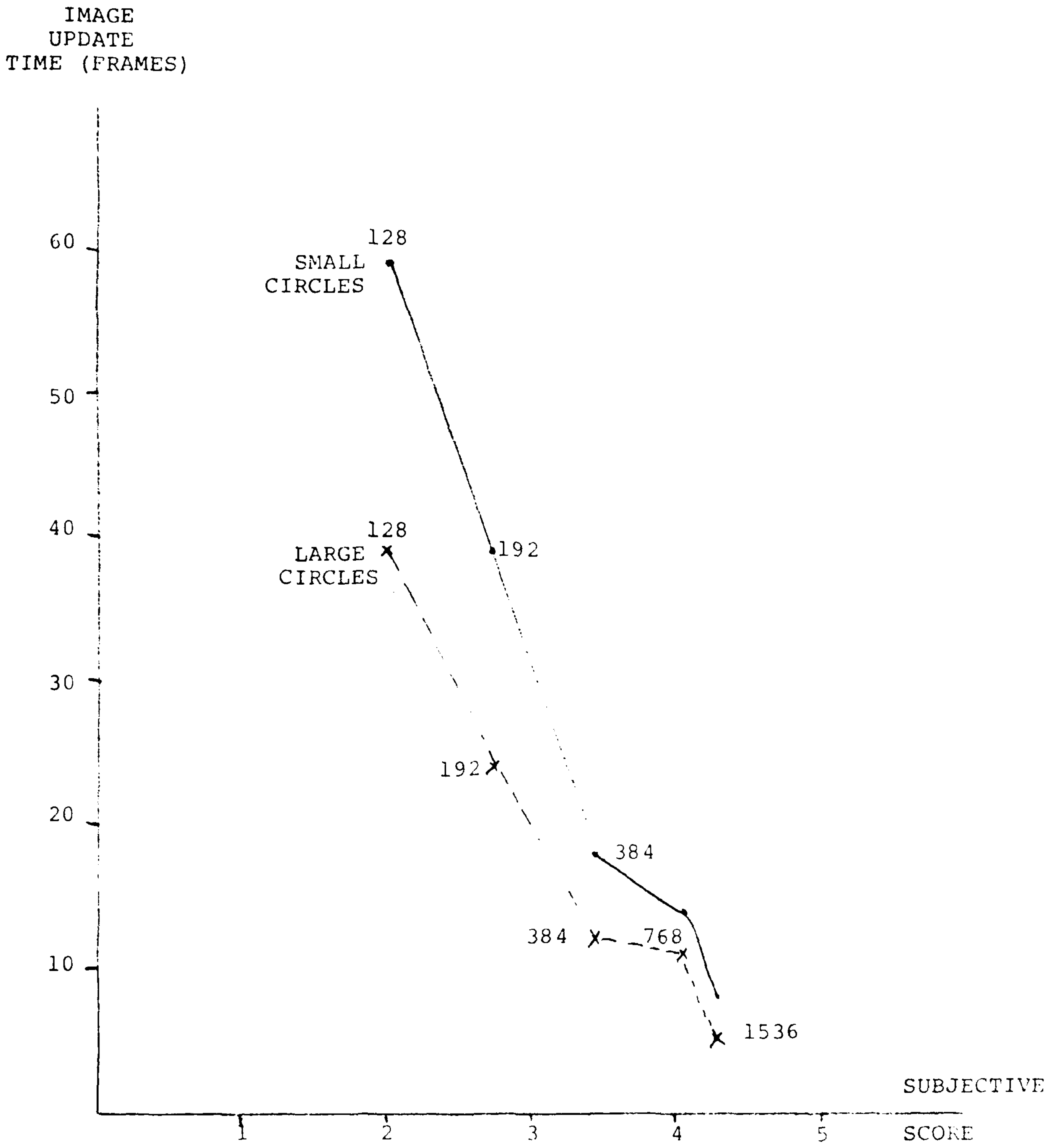


FIGURE 6.3 CORRELATION-SUBJECTIVE SCORE
vs. IMAGE UPDATE TIME

7.0 DISCUSSION

Not unexpectedly, this codec in accordance with CCITT Standard H.261 has many features different from several previously examined proprietary models. Being of more recent design, it incorporates many significant improvements and at the same time presents more challenges to the development of a meaningful and efficient test methodology. More experience will be needed to establish a complete end-to-end system performance and test standard.

A unique feature of this codec is its ability to provide two formats producing different quality pictures. H.261 specifies that all units must handle QCIF while only some need to provide CIF. The results of the tests described herein make the wisdom of this requirement questionable. The low sampling rates of QCIF limit the achievable spatial resolution to such a low value that the mean subjective rating at both 128 and 192 Kbps is "annoying" which is at best marginal even for videophone use. This is in spite of the fact that the transmitted frame rate is considerable higher than at CIF, producing a much less jerky picture. Comparing the image update times gives the same result.

CIF at 128 Kbps also produces an "annoying" picture, in this case evidently caused by excessive jerkiness. At 192 Kbps, the subjective rating indicates acceptable performance for videophone even though the jerkiness is still very high, but other presently undefined motion parameters appear to make up for it. At 384 Kbps, the subjective rating is better than "slightly annoying", and higher bit rates produce only "not annoying" impairments. These improvements in subjective rating are clearly tied to reduction and finally complete elimination of jerkiness -- the transmitted frame rate reaches its normal value of 30 -- and possibly also to trivially short image update times.

The relatively high bit rates needed to give even marginally acceptable performance at the low end of the range are due to the fact that the actual video bit rate is 58 Kbps lower than the total bit rate. Improved designs are expected to contain a 16 Kbps audio channel. This will increase the available video bit rate by 40 Kbps and probably provide about the same performance at the next lower step in the P x 64 hierarchy.

There are some features which may or may not have been taken into account by the evaluators but are quite obvious to an expert observer. The jerkiness caused by zooming appears to be less objectionable than that caused by panning. This may be due to the specific method of motion compensation used in this codec. The block coding specified by H.261 causes slightly noticeable blocking only at low rate CIF, mainly right after switching, which is much less objectionable than with previously examined codecs also using block coding. At QCIF a thin line drawn with a red pen is reproduced as black. This is apparently due to the very low resolution available for the color difference components. Color rendition of larger areas is not impaired.

8.0 CONCLUSION AND RECOMMENDATIONS

The effort described in this report shows that methods previously developed for the test of digital video systems incorporating proprietary design codecs can also be applied to future global systems with codecs built in accordance with CCITT Standard H.261. Some features of the new design resulting in different performance are readily recognized by the objective tests. The emphasis to be put onto the various objective performance parameters is changed from pıevious tests. Overall performance is generally improved but only for the full CIF mode. The low sampling rate of the QCIF mode makes it at best marginally acceptable for videophone pictures of persons. Correlation between subjective and objective tests is much better than in previous tests.

Whenever the transmission bit rate of a codec is reduced or the demand on the codec is increased by the transmitted material, some sacrifice of performance must be made. This can be achieved by reduction of quantizing and encoding accuracy resulting in a lower quality transmitted picture within the same time period, or by using more time to process the same quality picture which results in a lower transmitted frame rate. Most codec manufacturers use some combination of both methods which in their opinion gives the best overall result. The codec under test in this program obviously depends almost exclusively on frame rate reduction. This reduces the relevance of temporal response as a parameter for performance evaluation and for correlation of objective and subjective test results. This presents no technical problem since both temporal response and transmitted frame rate are derived from the same objective test. However, applications of temporal response and its relation to transmitted frame rate should be further investigated.

The subjective results obtained in these tests were evaluated only for overall performance using all types of available test material. In practice, it is often desirable to evaluate a digital video system in terms of a specific application. All the subjective test sequences can be readily assigned to one or possibly several application groups. Thus separate scores for various applications can be obtained from the available test results.

The objective test results described herein were obtained with encoder and decoder connected back-to-back. It must be emphasized that this was done for convenience only and is by no means a requirement for the performance of the tests. It has been demonstrated that the use of test tapes with carefully designed computer generated patterns makes it possible to perform objective measurements when only the decoder output signal is available. Without this feature, it is impossible to make practical end-to-end performance tests on a digital video system.

The use of the computer generated test patterns on high quality video tapes is fully acceptable but not the ultimate ideal. Recently, test equipment has become available which measures selected performance parameters of H.261 codecs. Whenever the capabilities of these equipments can be expanded to include all pertinent parameters, the implementation of objective end-to-end tests on digital

video systems will be greatly simplified.

At the time of the effort described herein, only one codec model in accordance with CCITT H.261 Standard was available. Other manufacturers are in the process of putting their equipment on the market. Since in spite of the common standard there are many features which are not specified, it is important that measurements on other models be performed. The main purpose of that is not the compare quality but to verify the applicability of the objective test methodology. Some static tests using standard analog test equipment should also be performed. Subsequently, it will be interesting to make system tests between encoders and decoders of different manufacturers. Ultimately, it will be desirable to test between NTSC and PAL system standards even though that will require a more complex test arrangement.

APPENDIX A

DRAFT STANDARDS

AMERICAN NATIONAL STANDARD

FOR TELECOMMUNICATIONS -

DIGITAL TRANSPORT OF VIDEO

TELECONFERENCING / VIDEO

TELEPHONY SIGNALS -

SYSTEM M-NTSC ANALOG

INTERFACE SPECIFICATIONS AND

PERFORMANCE PARAMETERS

FOREWORD

(This foreword is not part of the American National Standard).

This standard addresses the performance characteristics of **Video Teleconferencing/Video Telephony (VTC/VT)** service channels employing digital transport. The performance parameters are expressed as a function of **a single coding** between the end points of a transmission service channel. Each transmission service channel may be used by itself or may become part of a larger transmission channel. Performance definitions and measurement methods are provided if appropriate. Interface definitions are provided to facilitate compatibility between end users, service providers, and carriers.

This standard is intended to provide a means of assuring and maintaining transmission performance quality between defined transmission service channel interfaces.

Video signals originated or transmitted in accordance with standards other than system M-NTSC may not necessarily be compatible with the specifications of this standard.

Suggestions for improvement of this standard are welcome. They should be sent to The Exchange Carriers Standards Association - Committee T1 Secretariat, Suite 200, 5430 Grosvenor Lane, Bethesda, MD 20814-4505.

This standard was processed and approved for submittal to ANSI by Accredited Standards Committee T1 - Telecommunications. Committee approval of this standard does not necessarily imply that all committee members voted for its approval. At the time it approved this standard, Committee T1 had the following officers and members:

I. N. Knight, Chairman
A. K. Reilly, Vice-Chairman
O. J. Gusella, Jr., Secretary

Organization Represented	*Name of Representative*
EXCHANGE CARRIERS	
Alltel Service Corporation	
Ameritech Services, Inc.	
Bell Atlantic	
Bell Communications Research	
BellSouth Services	
Centel Corporation	
Contel Service Corporation	
Exchange Carriers Standards Association	
GTE SC/Telephone Operations	
National Telephone Cooperative Association	
NYNEX	
Pacific Bell	
Pureto Rico Telephone Company	
Southern New England Telephone	
Southwestern Bell Telephone Company	
Telephone & Data Systems	
United States Telephone Association	
United Telephone System, Inc.	
U S WEST	
INTEREXCHANGE CARRIERS	
AT&T Communications	
CNCP Telecommunications	
Communications Transmission, Inc.	
COMSAT	
Contel ASC	
MCI Telecommunications Corp.	
Siemens Communication Systems, Inc.	
Stromberg Carlson Corporation	
Telco Systems, Inc.	
Telecom Canada	
Telecommunications Techniques Corporation	
Telex Computer Products, Inc.	
TELINQ Systems, Inc.	
TIMEPLEX, Inc.	
Tymnet, Inc./McDonnell Douglas Corporation	
Unisys	
US Sprint	
U. S. Telecommunications Suppliers Association	
Verilink Corporation	
Video Conferencing Systems, Inc.	
Western Union Telegraph Company	

MANUFACTURERS

ADC Telecommunications, Inc. ....
Alcatel, NA ....
AMP Inc. ....
AT&T Technologies ....
Digital Equipment Corporation ....
DSC Communications Corporation ....
ECI Telecom Inc. ....
Ericsson, Inc. ....
Fortel Corporation ....
Fujitsu America, Inc. ....
General DataComm Industries ....
GTE Communication Systems Corporation ....
Harris Corporation ....
Hekimian Laboratories, Inc. ....
Hewlett Packard ....
Honeywell Bull ....
IBM Corporation ....
Integrated Technology, Inc. ....
LICOM, Inc. ....
M/A COM Corporation ....
Mitel Corporation ....
Motorola, Inc. ....
Myriad Concepts, Inc. ....
NCR Corporation ....
NEC America, Inc. ....
Northern Telecom Inc. ....
Phoenix Microsystems, Inc. ....
Plantronics, Inc. ....
Racal Milgo, Inc. ....
Rockwell International Corporation ....

GENERAL INTEREST

American Broadcasting Company ....
Ashford Associates ....
Association of Telemessaging Services International, Inc. ....
BDM Corporation, The ....
Carter Hawley Hale Information Services ....
CBS Broadcasting Group CBS, Inc. ....
Creative Communications Consulting ....
Defense Communications Agency ....
Federal Express Corporation ....
General Services Administration ....
Martin Marietta Information and Communications Systems ....
National Broadcasting Company ....
National Bureau of Standards ....

National Communications System ....
National Telecommunications and Information Administration,
Institute for Telecommunication Sciences (NTIA/ITS) ....
OMNICOM, Inc. ....
Rural Electrification Administration ....
Utilities Telecommunications Council ....

At the time it approved this standard, the Technical Subcommittee for Performance T1Q1 had the following officers and members:

Godfrey Williams, Chair
R. E. Jensen, Vice-Chair
R. B. Waller, Secretary

Work Group T1Q1.5 Wideband Program, which developed this standard had the following officers and participants:

Howard Meiseles, Chair
Steve Gloeckle, Vice-Chair
Warner Johnston, Secretary

List of members will be added later.

Participants

John Bekerich	David Bennett	Lyle Bradt
Ralph Brainard	Fred Camarillo	Hugh Chen
Gerard Clement	Bill Coufal	Tom Garcia
Steve Gloeckle	Ronald Gnidziejko	David Hanna
Hal Holton	Takunori Ito	James Jamison
Granger Kelley	Maurice Peronnet	Joe Presti
Edmund Quincy	Richard Quinn	Gary Rekstad
Michael Rossman	John Roth	Richard Schaphorst
Carl Sederquist	Alice Smith	Marvin Straus
Elmer Swope	C. F. Taylor	Bob Traill
Evert Turvey	Ed Underwood	Daniel Wirth
M. N. Woinsky	Stephen Wolf	William Zon

Editorial Committee of the Video Teleconferencing/Video Telephony Standard

Daniel Wirth, Chair
Gerard Clement
Bill Coufal
Granger Kelley
Edmund Quincy
Richard Quinn
Richard Schaphorst
Marvin Straus
C. F. Taylor
Stephen Wolf

TABLE OF CONTENTS

List of Figures

List of Tables

American National Standard for Telecommunications -

Digital Transport of Video Teleconferencing / Video Telephony Signals -

System M-NTSC Analog Interface Specifications and Performance Parameters

1 Scope, Purpose, and Application.

1.1 Scope.

This standard covers analog interface specifications in the system M-NTSC format and performance parameters of Video Teleconferencing/Video Telephony (VTC/VT) transmission service channels employing digital transport. VTC/VT signals created or transmitted in accordance with other standards or formats may not necessarily be compatible with the specifications of this standard.

This standard specifies the performance of transmission service channels employing digital transport provided to convey VTC/VT signals and their associated audio signals only. Performance values are for a single coding (refer to Figure 1) and are allocated by grades of service. Performance definitions and measurement methods are provided if appropriate. Interface specifications are provided to facilitate compatibility between end users, service providers, and carriers.

The performance characteristics identified within this standard apply to the transmission quality between the defined interfaces. Those interfaces are between VTC/VT transmission service providers and end users. This standard defines neither the interconnection nor the performance characteristics of specific apparatus or equipment.

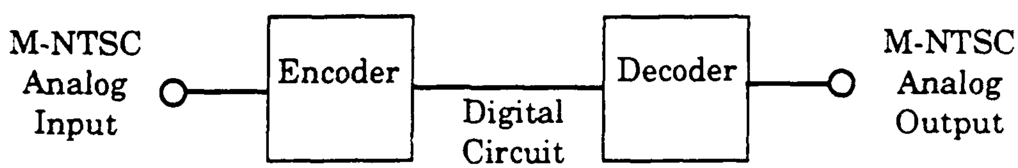


Figure 1 Digital Transmission Service Channel with Interfaces

1.2 Purpose.

The purpose of this standard is to assure the uniform application of standard values of transmission parameters for VTC/VT signals transported digitally by portions of the telecommunications network. It is intended to provide a common understanding by manufacturers, carriers, and their customers.

1.3 Application.

The primary applications of this standard are for specifying and evaluating the performance of a VTC/VT transmission service employing digital transport provided by common carriers. This service is used to transport the audio and video portions of the VTC/VT signals.

2 References and Related Standards.

2.1 American National Standards.

This standard is intended to be used in conjunction with the following American National Standards.

2.2 Other Related National Standards.

2.3 Other Related Standards.

2.4 Other Related Publications.

3 Definitions.

3.1 Special Word/Phrase Usage.

average picture level (APL): The average level of the picture signal during active scanning time integrated over a frame period and defined as a percentage of the range between blanking and reference white.

blocking: Distortion in the received video imagery characterized by rectangular or checkerboard patterns that were not present in the original video imagery before transmission.

blurring/smearing: Distortion in the received video imagery characterized by reduced sharpness of edges and detail that were present in the original video imagery before transmission.

coding: The digital encoding of an analog signal and decoding to an analog signal.

digital transport: A portion of the telecommunication network using digital methods for the transmission of signals from one point to another to complete a transmission service channel. A transmission service channel may have one or more

digital transport portion(s).

edge busyness: The deterioration of motion video such that the outlines of moving objects are displayed with randomly varying activity.

frame cuts: Video imagery where adjacent frames are not highly correlated.

image persistence: The appearance of earlier faded video frames of a moving and/or changing object within the current frame. (e.g., an object that was erased continues to appear in the received video imagery.)

jerkiness: The original smooth and continuous motion is perceived as a series of distinct 'snapshots'.

longitudinally balanced: A circuit is perfectly longitudinally balanced if disturbing common mode voltages (V_c) result in zero differential voltage (V_d). For circuits that are not perfectly longitudinally balanced, the degree of longitudinal balance is reported in dB using:

$$\text{Longitudinal Balance} = 20 \cdot \log_{10}\left(\frac{V_c}{V_d}\right)$$

mosquito noise: The noise generated by the block processing of moving objects that gives the appearance of false small moving objects (e.g., a mosquito flying around a person's head and shoulders).

motion response degradation: The deterioration of motion video such that the received video imagery has suffered a loss of spatio-temporal resolution.

motion video: Video imagery that conveys movement.

spatial application: For spatial applications, emphasis is placed on attaining high spatial resolution, possibly at the expense of reduced temporal positioning accuracy (or increased jerkiness). This application group is concerned with the ability to read small characters and see fine detail in still video and/or motion video which contains a very limited amount of motion.

spatial performance: A measure of the ability of a video transmission system to accurately reproduce still scenes.

still video: Video imagery that does not convey movement.

temporal application: For temporal applications, emphasis is placed on temporal positioning accuracy (or reduced jerkiness), possibly at the expense of reduced spatial resolution. This application group is concerned with the ability to accurately distinguish such items as facial expressions and lip movements in face to face and/or conference room settings.

temporal performance: A measure of the ability of a video transmission system to

accurately reproduce moving scenes.

transmission service channel: A transmission service channel is the one-way transmission path between two designated points (analog in, analog out).

video: (1) The visually displayed images of video teleconferencing/video telephony. (2) Of or pertaining to the visually displayed images of video teleconferencing/video telephony.

video frame: A single frame of video composed of two interlaced fields.

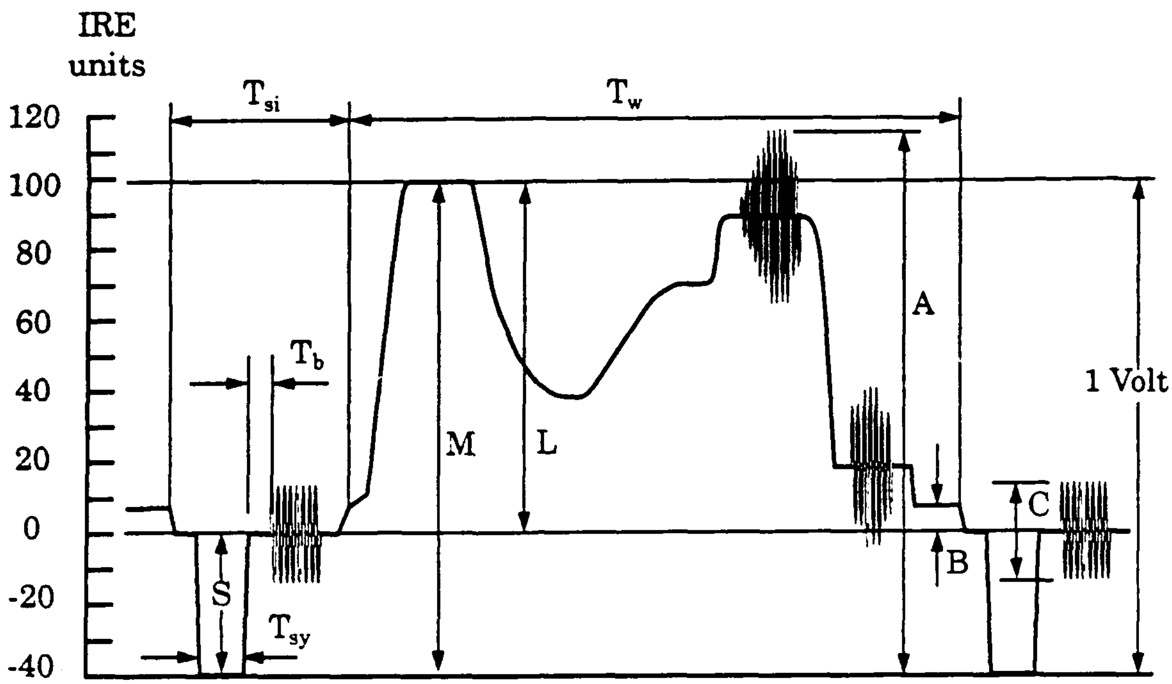
video imagery: A contiguous sequence of video frames.

video teleconferencing/video telephony motion artifacts: In a video teleconferencing/video telephony system, deteriorations of motion video due to image data compression that are observable by the viewer.

video teleconferencing/video telephony service (VTC/VT): The transmission of video signals capable of portraying motion and the accompanying audio signal(s) between two or more locations using digital transmission facilities. A typical example of this service is video teleconferencing between groups of personnel located at two or more locations.

3.2 Video Signals.

3.2.1 Video Signal Description. The waveform terminology used throughout the standard is in accordance with Figure 2, where the standard video signal waveform terminology is shown and measured in IRE units as shown in Figure 3.

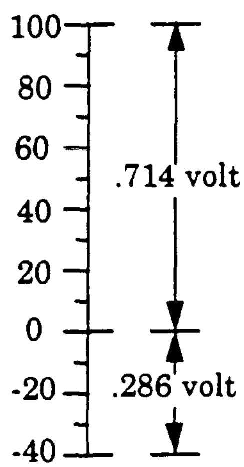


Waveform Terminology

A: The peak-to-peak amplitude of the composite color video signal
B: The difference between black level and blanking level (set-up)
C: The peak-to-peak amplitude of the color burst
L: Luminance signal - nominal value
M: Monochrome video signal peak-to-peak amplitude (M=L+S)
S: Synchronizing signal amplitude

T_b: Duration of breezeway
T_{si}: Duration of line blanking period
T_{sy}: Duration of line synchronizing pulse
T_w: Duration of active line period

Figure 2 Standard Video Signal General Waveform Terminology



(For a 1 V P-P composite signal)

Figure 3 IRE Unit Scale - Video

3.2.2 Test Signal Description. Reference to time (T) in the description of the following test signals refers to the half-amplitude pulse-width duration and not the rise time nor the fall time of a pulse transition and has a value of 125 nanoseconds.

3.2.3 Static Test Signals.

3.2.3.1 Composite Test Signal. The composite test signal shown in Figure 4 consists of a line bar (125 nanosecond rise time and fall time), a 2T pulse (250 nanosecond half-amplitude duration), a 12.5T (1.5625 microseconds half-amplitude duration) chrominance pulse, and a 5-riser staircase signal modulated by the color sub-carrier having a peak-to-peak amplitude of 40 IRE units superimposed upon standard synchronizing and blanking signals. Reference A and B are the measurement points utilized in the measurement of insertion gain and insertion-gain variation.

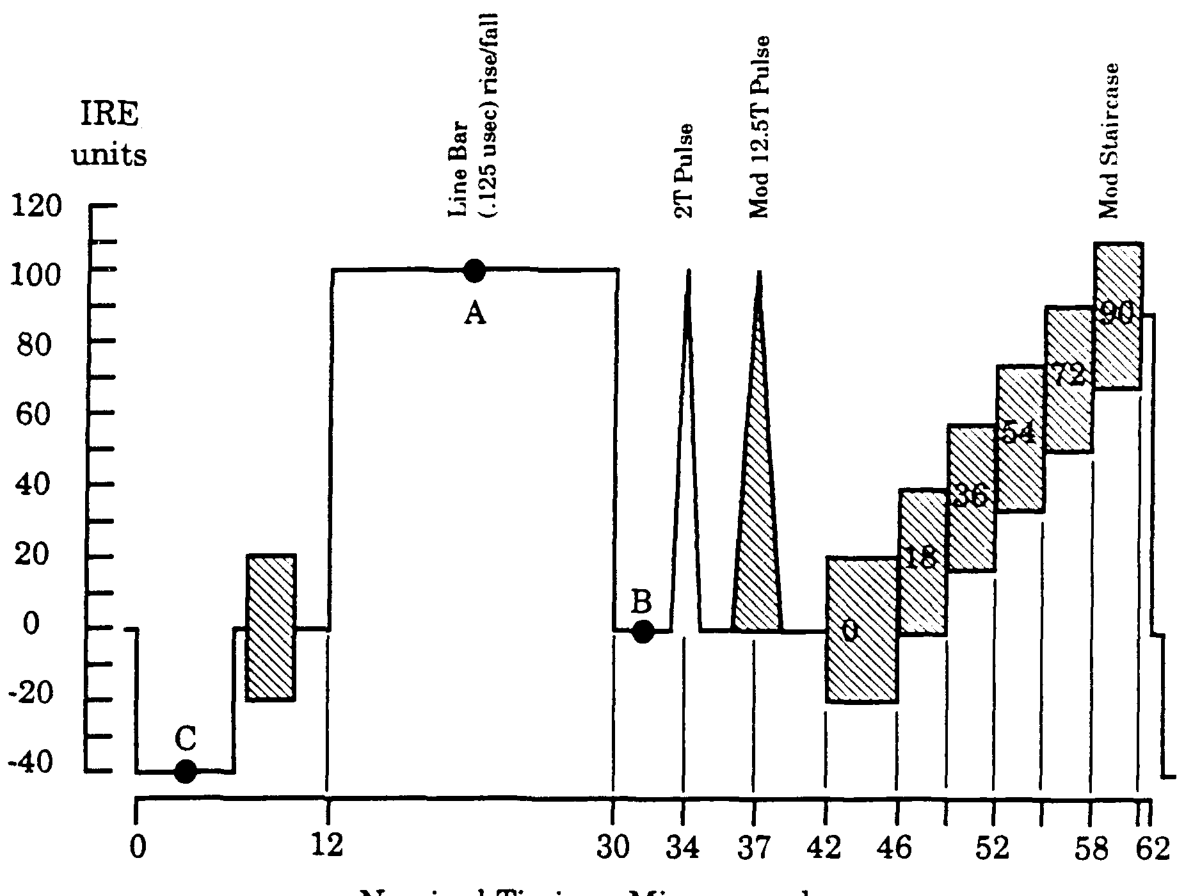


Figure 4 Composite Test Signal - Video

4 Baseband VTC/VT Interface Specifications.

4.1 Video Signal Electrical Interface Specifications.

4.1.1 Impedance.

4.1.1.1 Source Impedance.

4.1.1.1.1 Definition. The video source impedance of a transmission service channel, Z_s shown in Figure 5, is the impedance presented to the input terminals of a transmission service channel or other video baseband input point by the output terminals of the signal source. Proper source impedance is required for service channel evaluation.

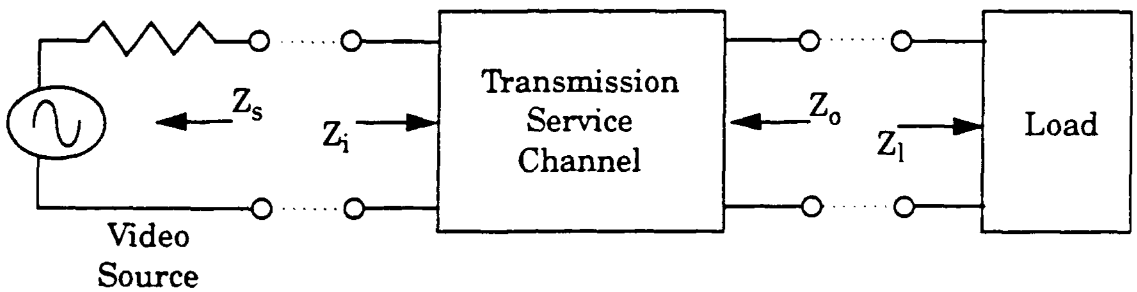


Figure 5 Impedance Reference - Unbalanced to Ground - Video

4.1.1.1.2 Standard Value. The standard value shall be 75 ohms, unbalanced to ground, with a return loss of at least x dB over the frequency range of 20 Hz - 4.2 MHz.

4.1.1.1.3 Method of Measurement. The impedance is measured by using impedance measurement equipment and the return loss is reported in dB using the following formula:

$$\text{Return Loss} = 20 \cdot \log_{10}\left|\frac{Z + Z_m}{Z - Z_m}\right|$$

Where

Z = specified standard impedance

Z_m = measured impedance

Alternately, the return loss may be measured using a return loss bridge.

4.1.1.2 Input Impedance.

4.1.1.2.1 Definition. The video input impedance of a transmission service channel, Z_i shown in Figure 5, is the impedance presented by the input terminals of a transmission service channel or other video baseband input point.

4.1.1.2.2 Standard Value. The standard value shall be 75 ohms, unbalanced to ground, with a return loss of at least x dB over the frequency range of 20 Hz - 4.2 MHz.

4.1.1.2.3 Method of Measurement. See section 4.1.1.1.3 on page 7.

4.1.1.3 Output Impedance.

4.1.1.3.1 Definition. The video output impedance of a transmission service channel, Z_0 shown in Figure 5, is the impedance presented by the output terminals of a transmission service channel or other baseband output point.

4.1.1.3.2 Value. The standard value shall be 75 ohms, unbalanced to ground, with a return loss of at least x dB over the frequency range of 20 Hz - 4.2 MHz.

4.1.1.3.3 Method of Measurement. See section 4.1.1.1.3 on page 7.

4.1.1.4 Load Impedance.

4.1.1.4.1 Definition. The video load impedance of a transmission service channel, Z_l shown in Figure 5, is the impedance presented by the input terminals of the device which will terminate the video baseband output of the transmission service channel. Proper load impedance is required for service channel evaluation.

4.1.1.4.2 Standard Value. The standard value shall be 75 ohms, unbalanced to ground, with a return loss of at least x dB over the frequency range of 20 Hz - 4.2 MHz.

4.1.1.4.3 Method of Measurement. See section 4.1.1.1.3 on page 7.

4.2 Video Signal.

4.2.1 Polarity of the Picture Signal.

4.2.1.1 Definition. The polarity of the picture signal of a transmission service channel is the sense of the potential of a portion of the signal representing a dark area of a scene relative to the potential of a portion of the signal representing a light area. Polarity is stated as "black positive" or "black negative". It is the polarity presented to the transmission service channel input terminals and presented by the transmission service channel output terminals.

4.2.1.2 Standard Value. The polarity of the picture signal is "black negative".

4.2.1.3 Method of Measurement. The polarity of the picture signal is

determined by use of an oscilloscope or waveform monitor of known deflection polarity.

4.2.2 Input Signal.

4.2.2.1 Signal Level.

4.2.2.1.1 Definition. The input signal level of a transmission service channel is the difference in voltage between sync tip (-40 IRE units) and reference white (100 IRE units) of a composite video signal presented to the video baseband input terminals. It is expressed in volts.

4.2.2.1.2 Standard Value. The standard value shall be a nominal 1 volt peak-to-peak difference between sync tip and reference white, 140 IRE units (see Figure 2 on page 5 and Figure 3 on page 5).

4.2.2.1.3 Method of Measurement. The input signal level is measured by means of a properly calibrated and terminated oscilloscope or waveform monitor.

4.2.2.2 Time Base Error.

4.2.2.2.1 Definition. Time base error is defined as the difference between the instantaneous time base (which is the time between the 50% value of the leading edges of two successive horizontal sync pulses) and the time base averaged over one frame. The time base error defined here is typically that which is generated by a video tape machine.

4.2.2.2.2 Standard Value. The absolute value of the largest time base error shall be less than x microseconds.

4.2.2.2.3 Method of Measurement. The time base error is measured by applying the video signal to a time base error reading instrument.

4.2.2.3 Unweighted Signal to Noise Ratio.

4.2.2.3.1 Definition. The unweighted signal-to-noise ratio of the input signal is the ratio of the peak-to-peak luminance signal, blanking to reference white (nominally 0.714 volt = 100 IRE units), to the rms band limited noise level. The noise may be mixed random and quantizing noise. Synchronizing signals are not included in the signal measurement. It is measured at the output of the video source.

4.2.2.3.2 Standard Value. The standard value is greater than x dB. In addition, the input of the transmission service channel must tolerate noise that is out of band and could have voltage spikes several times that of the band limited noise.

4.2.2.3.3 Method of Measurement. The input signal of a transmission service channel is connected to the input of a rms reading meter through standard low pass filter x and high pass filter y. The video S/N is measured using a video line that

is at a constant IRE level. The readings are averaged over a .4 second interval where synchronizing signals are excluded. This measurement is equivalent to the unweighted version of the signal to noise ratio measurement commonly known as NTC 7 3.16.

4.2.3 Output Signal.

4.2.3.1 Signal Level.

4.2.3.1.1 Definition. The output signal level of a transmission service channel is the difference in voltage between sync tip and reference white of a composite test signal presented by the video baseband output terminals. It is expressed in volts peak-to-peak.

4.2.3.1.2 Standard Value. The standard value shall be a nominal 1 volt peak-to-peak difference between sync tip and reference white, 140 IRE units (see Figure 2 on page 5 and Figure 3 on page 5).

4.2.3.1.3 Method of Measurement. Apply the composite test signal shown in Figure 4 on page 6 to the transmission service channel input. The output signal level is measured by means of a properly calibrated and terminated oscilloscope or waveform monitor.

4.2.3.2 Time Base Error.

4.2.3.2.1 Definition. Time base error is defined as the difference between the instantaneous time base (which is the time between the 50% value of the leading edges of two successive horizontal sync pulses) and the time base averaged over one frame.

4.2.3.2.2 Standard Value. The absolute value of the largest time base error shall be less than x microseconds.

4.2.3.2.3 Method of Measurement. See section 4.2.2.2.3 on page 9.

4.2.3.3 Non-Useful DC Component.

4.2.3.3.1 Definition. The non-useful DC component of the output signal is any DC component which is unrelated to the output signal. It will be present only as a result of the transmission equipment.

4.2.3.3.2 Standard Value. The standard value is ≤ x IRE units across a standard termination (75 ohms).

4.2.3.3.3 Method of Measurement. The non-useful DC component of the output signal is measured with a properly terminated DC coupled waveform monitor (WFM). Apply the composite test signal shown in Figure 4 on page 6 to the input terminals of the transmission service channel. With the WFM in a non-clamping

mode, measure the signal at the output of the transmission service channel. Terminate the transmission service channel input in 75 ohm and measure the new position of the trace. The non-useful DC component is the difference in IRE units between the 0 and new trace position.

4.2.4 Input to Output Video Relationships.

4.2.4.1 Active Video Area.

4.2.4.1.1 Definition. The active video area is the portion of the input video signal that is not blanked by the transmission service channel.

4.2.4.1.2 Standard Value. The active video area shall be from x_1 usec to x_2 usec of video lines y_1 through y_2 inclusive.

4.2.4.1.3 Method of Measurement. Inject a video signal into the input of the transmission service channel with a marker at each of the following 4 locations: 1.) line y_1, x_1 usec, 2.) line y_1, x_2 usec, 3.) line y_2, x_1 usec, 4.) line y_2, x_2 usec. Check the output of the transmission service channel for the presence of the markers. If all 4 markers are observed, then the entire active video area is being passed by the transmission service channel.

4.2.4.2 Active Video Shift.

4.2.4.2.1 Definition. The active video shift is the amount of vertical and horizontal shift of the output active video area with respect to the input.

4.2.4.2.2 Standard Value. The maximum shift for any of the 4 markers shall be y lines, x usec.

4.2.4.2.3 Method of Measurement. Using the test signal of section 4.2.4.1.3, measure the locations (line, usecs) of each of the 4 markers at the output of the transmission service channel. Compute the absolute value of the shift for each of the 4 markers by comparing the output marker locations to the input marker locations.

4.3 Audio Signal Electrical Interface Specifications.

4.3.1 Balanced Audio Specifications.

(The analog audio input to and output from the transmission service channel shall be balanced with respect to ground).

4.3.1.1 Input Common Mode Rejection Ratio (CMMR).

4.3.1.1.1 Definition. Input common mode rejection ratio indicates the degree to which unwanted signals coupled into both sides of a balanced line are rejected by the input of the transmission service channel.

4.3.1.1.2 Standard Value. The common mode rejection ratio for all audio input channels to the transmission service channel shall be greater than x dB within

the respective service bandpass and at least y dB at 60 Hz. This specification must be met for common voltages up to x volts rms.

4.3.1.1.3 Method of Measurement. The audio input to the transmission service channel is connected to a center-taped 600 ohm resistive termination. A common mode test signal with voltage V_c is applied between the center tap and ground. The voltage of the resulting signal at the differential output of the transmission service channel is then measured (V_0). The common mode rejection ratio is given by:

$$CMRR = 20 \cdot \log_{10}\left(\frac{V_c}{V_o}\right)$$

4.3.1.2 Output Common Mode Noise.

4.3.1.2.1 Definition. In a balanced audio system, any voltage that is common to both signal leads (measured with respect to ground) is common mode noise.

4.3.1.2.2 Standard Value. The common mode noise of all audio outputs from the transmission service channel shall be at least x dB below nominal operating level.

4.3.1.2.3 Method of Measurement. The audio output is terminated with a center-tapped 600 ohm load. The common mode noise is the voltage measured between the center tap of the 600 ohm load and ground.

4.3.2 Impedance.

4.3.2.1 Source Impedance.

4.3.2.1.1 Definition. The audio source impedance of a transmission service channel, Z_S, shown in Figure 6, is the impedance presented to the input terminals of a transmission service channel by the output terminals of the signal source. Proper source impedance is required for transmission service channel evaluation.

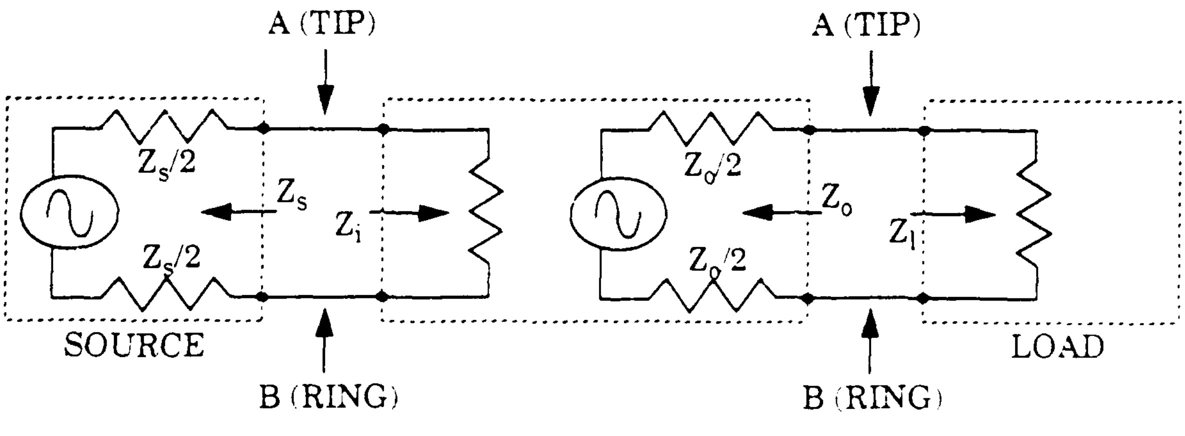


Figure 6 Impedance Reference - Balanced to Ground - Audio

4.3.2.1.2 Standard Value. The standard value shall be 600 ohms, nominally resistive, highly longitudinally balanced, and with a return loss of at least x dB over the frequency range of the respective service bandpass.

4.3.2.1.3 Method of Measurement. The impedance is measured by using impedance measurement equipment and the return loss is reported in dB using the following formula:

$$\text{Return Loss} = 20 \cdot \log_{10}\left|\frac{Z + Z_m}{Z - Z_m}\right|$$

Where

Z = specified standard impedance

Z_m = measured impedance

Alternately, the return loss may be measured using a return loss bridge.

4.3.2.2 Input Impedance.

4.3.2.2.1 Definition. The input impedance of a transmission service channel Z_i shown in Figure 6, is the impedance presented by the input terminals of a transmission service channel.

4.3.2.2.2 Standard Value. The standard value is 600 ohms, nominally resistive, highly longitudinally balanced, and with a return loss of at least x dB over the frequency range of the respective service bandpass.

4.3.2.2.3 Method of Measurement. See section 4.3.2.1.3 on page 13.

4.3.2.3 Output Impedance.

4.3.2.3.1 Definition. The output impedance of a transmission service channel, Z_0 shown in Figure 6, is the impedance presented by the output terminals of a transmission service channel.

4.3.2.3.2 Standard Value. The standard value is 600 ohms, nominally resistive, highly longitudinally balanced, and with a return loss of at least x dB over the frequency range of the respective service bandpass.

4.3.2.3.3 Method of Measurement. See section 4.3.2.1.3 on page 13.

4.3.2.4 Load Impedance.

4.3.2.4.1 Definition. The load impedance of a transmission service channel. Z_l shown in Figure 6, is the impedance presented by the input terminals of the device which will terminate the audio output of the transmission service channel. Proper load impedance is required for channel evaluation.

4.3.2.4.2 Standard Value. The standard value is 600 ohms, nominally resistive, highly longitudinally balanced, and with a return loss of at least x dB over the frequency range of the respective service bandpass.

4.3.2.4.3 Method of Measurement. See section 4.3.2.1.3 on page 13.

4.4 Audio Signal.

4.4.1 Input Signal.

4.4.1.1 Input Signal Level.

4.4.1.1.1 Definition. The input signal level to a transmission service channel is the signal level across the transmission service channel input impedance. When the signal is sinusoidal the input signal level is expressed in dBm.

4.4.1.1.2 Standard Value. The peak operating level of the input signal to the transmission service channel across the standard impedance is equal to the peaks of a sine-wave whose average power is 0 dBm. The nominal level is -16 dBm, and the minimum clip level is +9 dBm.

NOTE:

A. Clip level is defined as the level at which the total distortion exceeds 1%.

B. This 9 dBm level is only to be applied at the end users interface for program audio services as described in this document. This level is only used for an out-of-service test.

4.4.1.1.3 Method of Measurement. The audio input signal level is measured by properly terminated audio test equipment.

4.4.1.2 Non-Useful DC Component.

4.4.1.2.1 Definition. The non-useful DC component of the audio signal is any DC component which is unrelated to the audio signal and is also present when the audio signal is muted.

4.4.1.2.2 Loop Requirements. When the audio signal is muted, direct current flow shall be equal to or less than x mA through a zero ohm termination across the audio source output terminals.

4.4.1.2.3 Longitudinal Requirements. Direct current flow shall be equal to or less than x mA when the terminals are shorted together and the current flow is measured through a zero impedance to ground.

4.4.2 Output Signal.

4.4.2.1 Output Signal Level.

4.4.2.1.1 Definition. The output signal level of a transmission service channel is the signal level across the transmission service channel output impedance.

4.4.2.1.2 Standard Value. When a 1 KHz test tone at the nominal operating level of -16 dBm is applied to the input of the transmission service channel, the output level shall be -16 dBm plus or minus x dB.

4.4.2.1.3 Method of Measurement. See section 4.4.1.1.3 on page 14.

4.4.2.2 Non-Useful DC Component.

4.4.2.2.1 Definition. The non-useful DC component of the audio signal is any DC component which is unrelated to the audio signal and is also present when the audio signal is muted.

4.4.2.2.2 Loop Requirements. When the audio signal is muted, direct current flow shall be equal to or less than x mA through a zero ohm termination across the transmission service channel output terminals.

4.4.2.2.3 Longitudinal Requirements. Direct current flow shall be equal to or less than x mA when the terminals are shorted together and the current flow is measured through a zero impedance to ground.

4.4.3 Input to Output Audio Relationships.

4.4.3.1 Signal Polarity.

4.4.3.1.1 Definition. The polarity of the signal is the polarity sense of a nonsymmetrical audio transient signal on the A (tip) terminal with respect to the B (ring) terminal of the balanced pair (see Figure 6 on page 12).

4.4.3.1.2 Standard Value. The polarity sense of the audio signal at the output of the transmission service channel shall be the same as that at the input of the transmission service channel.

4.4.3.1.3 Method of Measurement. A nominal 400-Hz clipped sine wave (1/2 wave rectified, see Figure 7) is fed into the input of the transmission service channel. An oscilloscope of a known deflection is used at both ends of the transmission service channel to determine the polarity sense of the audio signal.

Figure 7 Nonsymmetrical Test Waveform - Audio

5 Out-of-Service Baseband VTC/VT Performance Specification.

This out-of-service VTC/VT performance specification gives the average performance of the transmission service channel for a representative group of VTC/VT video and audio signals. The user is cautioned that the performance for a particular input signal may vary from the average performance presented here. If the performance for a particular input signal is desired, the user of this standard should refer to section 6 on page 20. The block diagram for out-of-service testing is shown in Figure 8. As shown in the figure, a test signal and/or test scene generator is connected to the input of the transmission service channel. The test generator provides a source of known audio and video signals for testing the transmission service channel. Parameter measurement equipment is connected to the output of the transmission service channel. The provision is made for sending telemetric data from the test generator to the parameter measurement equipment. The exact requirements for the telemetric data have yet to be determined (the requirements may be as simple as a begin test command). Several options for transmitting the telemetric data include a separate data channel, and the use of the vertical blanking interval of the video signal (if transmitted by the transmission service channel). If the input and the output of the transmission service channel are co-located and non-automated testing is acceptable, then the telemetric data path may be unnecessary for some measurements. The parameter measurement equipment shown in Figure 8 determines the performance parameter values in this standard. Then, these parameter values are inserted into a set of performance calculations to obtain video and audio performance ratings for the transmission service channel. The final performance ratings may then be (optionally) transmitted back to the input of the transmission service channel.

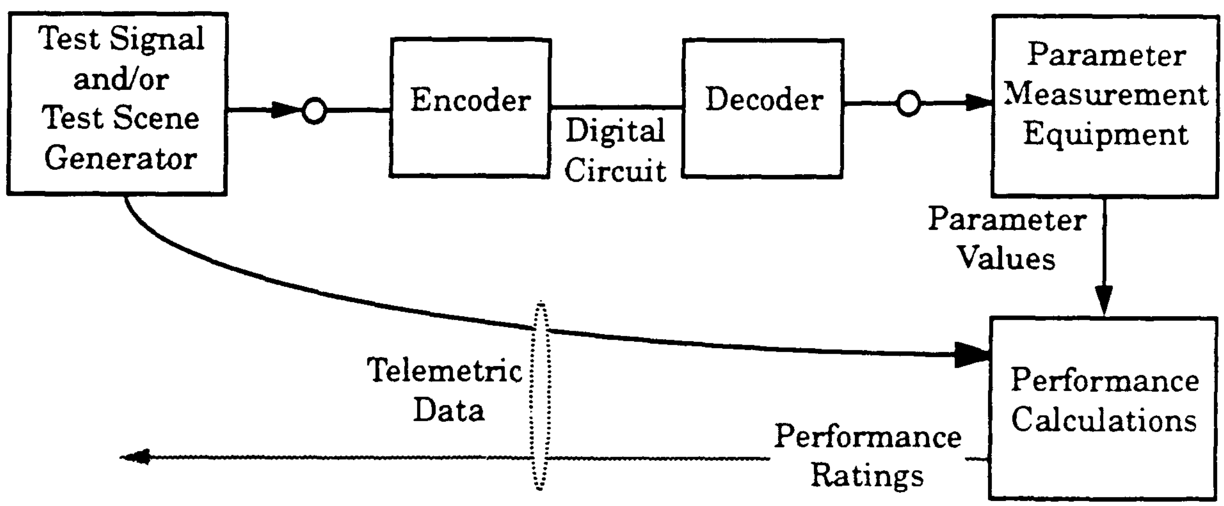


Figure 8 Out-of-Service Testing Block Diagram

5.1 Video Performance.

The video performance has been divided into two major areas; spatial

performance and temporal performance. Spatial performance, a measure of the ability of a video transmission system to accurately reproduce still scenes, is a primary concern for graphical or spatial applications (see definition on page 3). Temporal performance, a measure of the ability of a video transmission system to accurately reproduce moving scenes, is a primary concern for temporal applications (see definition on page 3). A set of performance parameters for quantifying the spatial performance of a transmission service channel is given in section 5.1.1. The overall spatial performance rating for the transmission service channel is then obtained by inserting these parameter values into the algorithm given in section 5.1.2. A set of performance parameters for quantifying the temporal performance of a transmission service channel is given in section 5.1.3. The overall temporal performance rating for the transmission service channel is then obtained by inserting these parameter values into the algorithm given in section 5.1.4. The video performance application table in section 5.1.5 on page 18 can be used to specify the required spatial and temporal performance of the transmission service channel for a number of purposes (e.g., videophone, videoconference, etc.) and subject material (e.g., talking head, graphics, etc.).

5.1.1 Spatial Performance Measures. (This is the place to insert the useful traditional analog parameters as well as newer digital parameters that measure spatial distortions).

5.1.1.1 Parameter Name.

5.1.1.1.1 Definition.

5.1.1.1.2 Method of Measurement.

5.1.2 Spatial Performance Calculation. To compute the overall spatial performance rating for the transmission service channel (O_S), the measured parameter values in section 5.1.1 are inserted into the following equation: (This equation has been included as an example algorithm only)

$$O_s = \sum_{i=1}^{N} (c_i \cdot p_i) + c_0$$

Where p_i is the value of spatial parameter i, N is the total number of spatial parameters, and c_i (for i=0 to N) are constants determined according to the methods specified in Appendix A (i.e., the spatial performance equation shall accurately predict the subjective quality of spatial test scenes).

5.1.3 Temporal Performance Measures. (This is the place to insert the newer digital parameters that measure motion artifacts and temporal distortions).

5.1.3.1 Parameter Name.

5.1.3.1.1 Definition.

5.1.3.1.2 Method of Measurement.

5.1.4 Temporal Performance Calculation. To obtain the overall temporal performance rating for the transmission service channel (O_t), the measured parameter values in section 5.1.3 are inserted into the following equation: (This equation has been included as an example algorithm only)

$$O_t = \sum_{i=1}^{N} (c_i \cdot p_i) + c_0$$

Where p_i is the value of temporal parameter i, N is the total number of temporal parameters, and c_i (for i=0 to N) are constants determined according to the methods specified in Appendix A (i.e., the temporal performance equation shall accurately predict the subjective quality of temporal test scenes).

5.1.5 Video Performance Application Table. This table is provided as a guide to assist end-users and service-providers in specifying spatial and temporal performance levels. The spatial performance calculation (from section 5.1.2 on page 17) and the temporal performance calculation (from section 5.1.4 on page 18) are utilized as shown in Table 1. In the table, spatial performance increases as one moves from left to right and temporal performance increases as one moves down. The recommended spatial and temporal performance levels for various purposes (e.g., videophone, videoconference, etc.) and subject material (e.g., talking head, graphics, etc.) are shown.

	Spatial Performance			
Temporal Performance	Level 1 ($O_s>x_1$)	Level 2 ($O_s>x_2$)	Level 3 ($O_s>x_3$)	Level 4 ($O_s>x_4$)
Level 1 ($O_t>y_1$)			Videophone (graphics)	
Level 2 ($O_t>y_2$)	Videophone (talking head)	Videoconference (people)	Videoconference (people + graphics)	
Level 3 ($O_t>y_3$)		VCR (varied)	Distribution (training, education)	Contribution (TV studio)

Table 1 Video Performance Specifications for Several Applications

5.2 Audio Performance.

A set of performance parameters for quantifying the audio performance of a transmission service channel is given in section 5.2.1. The overall audio performance rating for the transmission service channel is then obtained by inserting these parameter values into the algorithm given in section 5.2.2. The audio performance application table in section 5.2.3 on page 19 can be used to specify the required audio performance of the transmission service channel for a number of purposes (e.g., videophone, videoconference, etc.).

5.2.1 Audio Performance Measures. (This is the place to insert the useful traditional audio parameters as well as newer digital parameters that measure audio distortions, such as those being considered by the CCITT).

5.2.1.1 Parameter Name.

5.2.1.1.1 Definition.

5.2.1.1.2 Method of Measurement.

5.2.2 Audio Performance Calculation. To obtain the overall audio performance rating for the transmission service channel (O_a), the measured parameter values in section 5.2.1 are inserted into the following equation: (This equation has been included as an example algorithm only)

$$O_a = \sum_{i=1}^{N} (c_i \cdot p_i) + c_0$$

Where p_i is the value of audio parameter i, N is the total number of audio parameters, and c_i (for i=0 to N) are constants determined according to methods similar to those specified in Appendix A (i.e., the audio performance equation shall accurately predict the subjective quality of audio signals).

5.2.3 Audio Performance Application Table. This table is provided as a guide to assist end-users and service-providers in specifying audio performance levels. The audio performance calculation (from section 5.2.2 on page 19) is utilized as shown in Table 2. In the table, audio performance increases as one moves from left to right. The recommended audio performance level for various purposes (e.g., videophone, videoconference, etc.) is shown.

Audio Performance			
Level 1 ($O_a>x_1$)	Level 2 ($O_a>x_2$)	Level 3 ($O_a>x_3$)	Level 4 ($O_a>x_4$)
Videophone	Videoconference	VCR, Distribution	Contribution

Table 2 Audio Performance Specifications for Several Applications

5.3 Audio-Visual Performance Measures.

Audio-visual performance measures quantify important attributes of the audio-visual signal. These performance measures fall into one of two categories: (1) parameters that quantify important attributes of both the audio signal and the video signal (2) parameters that quantify the interactions between the audio signal and the video signal, and hence require both the audio signal and the video signal for proper measurement.

5.3.1 Overall Path Delay.

5.3.1.1 Definition. Overall path delay is the greater of the one-way transmission delay of the audio signal, and the one-way transmission delay of the video signal, when only lips and eyes of the talking user (talking head) are moving.

5.3.1.2 Standard Value. (Under Study)

5.3.1.3 Method of Measurement. (Under Study)

5.3.2 Audio-Visual Synchronization.

5.3.2.1 Definition. Audio-visual synchronization is the difference between the one-way transmission delays of the audio signal and the video signal when only lips and eyes of the talking user (talking head) are moving.

5.3.2.2 Standard Value. (Under Study)

5.3.2.3 Method of Measurement. (Under Study)

6 In-Service Baseband VTC/VT Performance Specification.

In general, the performance of the transmission service channel depends upon

the information content of the input video and audio signals. This in-service VTC/VT performance specification provides a non-intrusive method for measuring the performance of the transmission service channel for any input signal. The block diagram for in-service testing is shown in Figure ~~8~~ 9. As shown in the figure, parameter measurement equipment is connected to both the input and output of the transmission service channel. The parameter measurement equipment is connected in a non-intrusive manner so that the video and audio performance of the transmission service channel are not effected. The provision is made for sending telemetric data from the parameter measurement equipment at the input to the parameter measurement equipment at the output. The exact requirements for the telemetric data have yet to be determined but will likely consist of extracted parameter values from the input signal. Several options for transmitting the telemetric data include a separate data channel, and the use of the vertical blanking interval of the video signal (if transmitted by the transmission service channel). The parameter measurement equipment determines the performance parameter values in this standard. Then, these parameter values are inserted into a set of performance calculations to obtain video and audio performance ratings for the transmission service channel. The final performance ratings may then be (optionally) transmitted back to the input of the transmission service channel. Performance ratings are calculated for each x second time interval of the transmission.

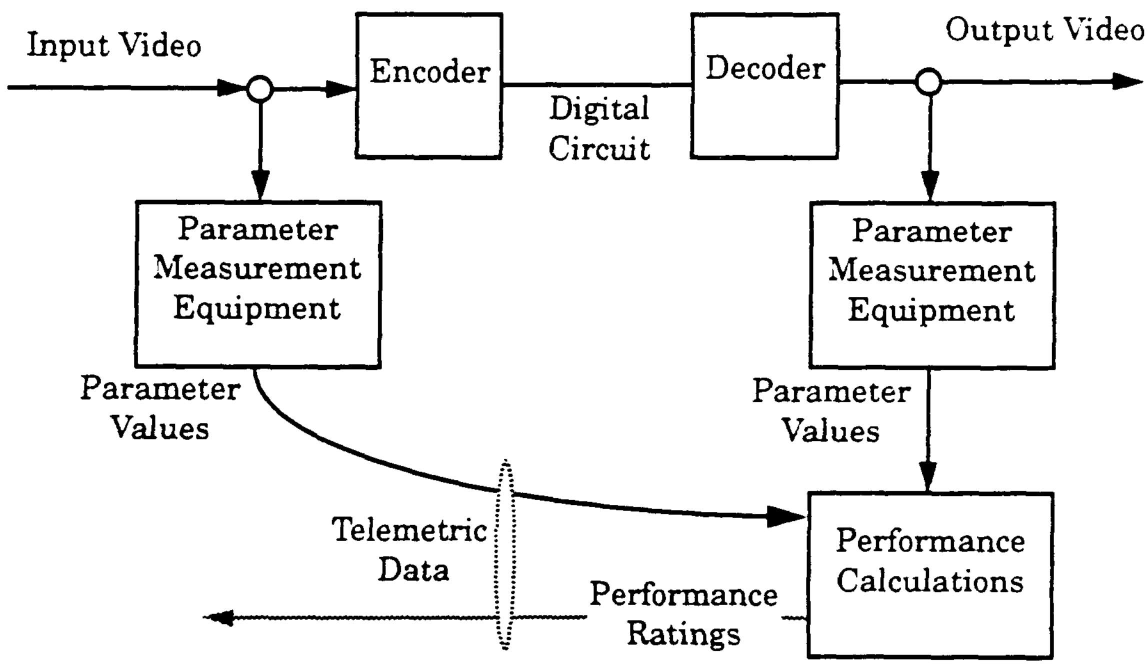


Figure 9 In-Service Testing Block Diagram

6.1 Video Performance.

A set of performance parameters for quantifying the video performance of a

transmission service channel is given in section 6.1.1. The overall video performance rating for the transmission service channel is then obtained by inserting these parameter values into the algorithm given in section 6.1.2.

6.1.1 Video Performance Measures.

6.1.1.1 Parameter Name.

6.1.1.1.1 Definition.

6.1.1.1.2 Method of Measurement.

6.1.2 Video Performance Calculation. To compute the overall video performance rating for the transmission service channel (O_v), the measured parameter values in section 6.1.1 are inserted into the following equation: (This equation has been included as an example algorithm only)

$$O_v = \sum_{i=1}^{N} (c_i \cdot p_i) + c_0$$

Where p_i is the value of video parameter i, N is the total number of video parameters, and c_i (for i=0 to N) are constants determined according to the methods specified in Appendix A (i.e., the video performance equation shall accurately predict the subjective quality of video test scenes).

6.2 Audio Performance.

A set of performance parameters for quantifying the audio performance of a transmission service channel is given in section 6.2.1. The overall audio performance rating for the transmission service channel is then obtained by inserting these parameter values into the algorithm given in section 6.2.2.

6.2.1 Audio Performance Measures.

6.2.1.1 Parameter Name.

6.2.1.1.1 Definition.

6.2.1.1.2 Method of Measurement.

6.2.2 Audio Performance Calculation. To obtain the overall audio performance rating for the transmission service channel (O_a), the measured parameter values in section 5.2.1 are inserted into the following equation: (This equation has been included as an example algorithm only)

$$O_a = \sum_{i=1}^{N} (c_i \cdot p_i) + c_0$$

Where p_i is the value of audio parameter i, N is the total number of audio parameters, and c_i (for i=0 to N) are constants determined according to methods similar to those

specified in Appendix A (i.e., the audio performance equation shall accurately predict the subjective quality of audio signals).

6.3 Audio-Visual Performance Measures.

6.3.1 Availability of Service. (Under Study)

Appendix

Appendix A

Methodology Used for Evaluating Out-of-Service Performance Measures

The information in this appendix provides a historical record of the approach and the actual test results that were used to select the objective performance measures in this standard. The detailed methodology presented here can by used by others to verify the validity of this standard as well as to expand upon the standard so that it will be useful for measuring the performance of future, as of now unspecified, audio-visual services. This appendix describes the methodology used for developing the video performance measures. Extension of this methodology to include the audio performance measures should be straightforward. The extended methodology should consider the possibility of interactions between the audio performance and the video performance.

Overview of Approach. An overview of the approach that is used to develop and evaluate the objective quality estimators is shown in Figure A1 on page 26. A primary design goal for objective quality estimators is an output that accurately predicts the overall quality score as perceived by the end-user. Thus, subjective test results form a key criterion for the evaluation of acceptable objective quality estimators. The test scenes used for the subjective tests are divided into two groups; spatial test scenes (s_1, s_2, ..., s_m), and temporal test scenes (t_1, t_2, ..., t_n). The spatial test scenes are chosen to test the spatial performance of the digital transmission service channel under test (DUT), while the temporal test scenes are chosen to test the temporal performance of the DUT. These subjective test scenes are input into a set of representative DUTs and the resulting output scenes are subjectively judged by viewing panels. Subjective impairment scores for these test scenes are then combined with the subjective combiners to produce two composite subjective scores for each DUT; a composite subjective score for the set of spatial test scenes (S_s), and a composite subjective score for the set of temporal scenes (S_t). Similarly, the objective test waveforms (w_1, w_2, ..., w_j) are also input into the DUTs and the proposed objective quality estimators (1 through k) each measure a set of objective parameters and produce objective quality estimates of the spatial performance (O_{s1}, O_{s2}, ..., O_{sk}) and temporal performance (O_{t1}, O_{t2}, ..., O_{tk}) of the DUT. The criterion for determining the accuracy of the objective quality estimates is minimization of the squared error differences between the objective and subjective estimates of performance averaged over all DUTs. These squared error differences are denoted by E_s^2 and E_t^2 in Figure A1, where E_s is the error in the spatial performance estimate and E_t is the error in the temporal performance estimate for a DUT. All of the outputs from the objective quality estimators can themselves be combined by an objective combiner. The final composite objective estimates from the objective combiner may provide improved estimates of the spatial performance (O_s) and temporal performance (O_t) of the DUT. This approach allows different organizations to independently develop objective

quality estimators and enables the integration of the various methods into a coherent standard.

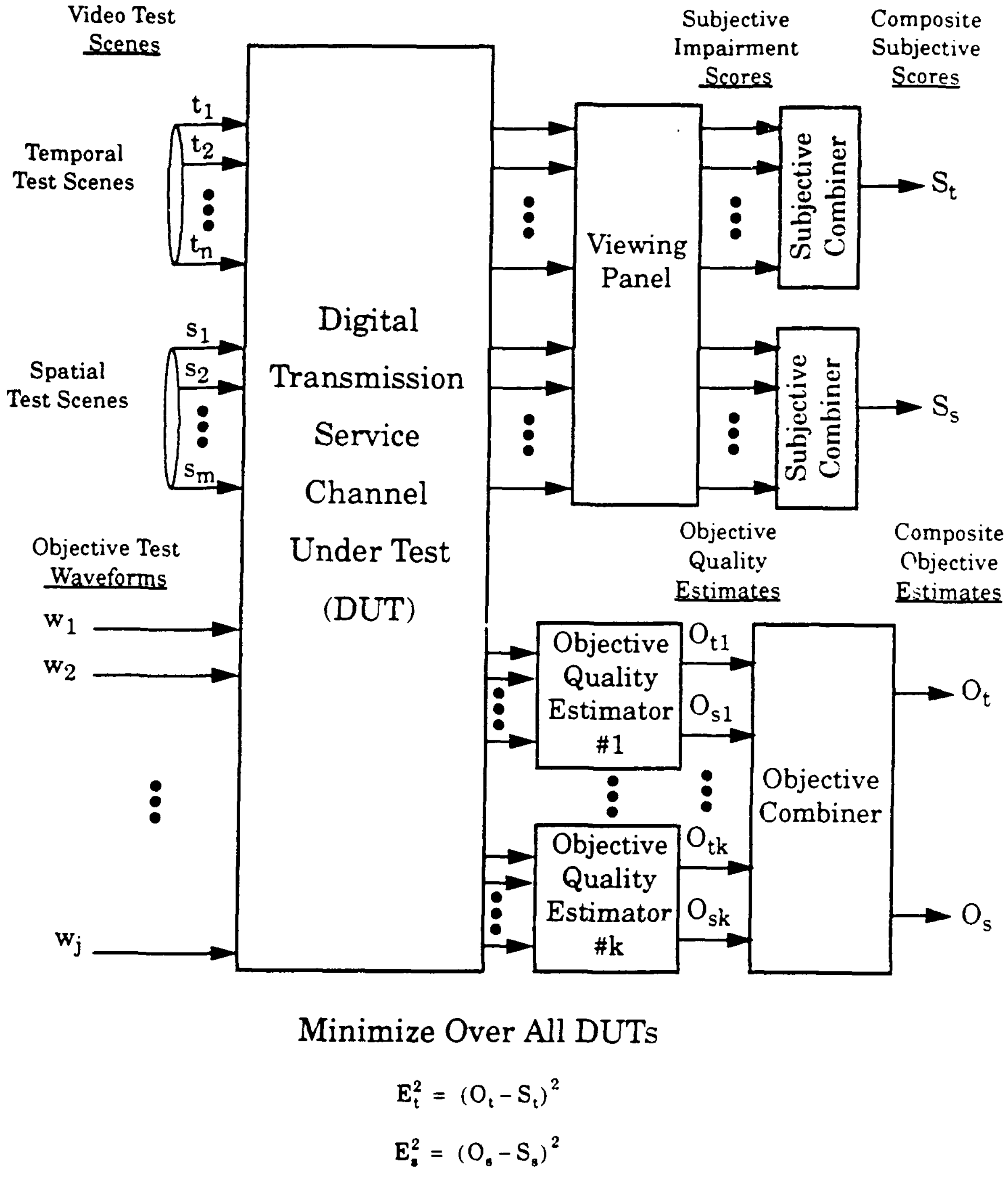


Figure A1 Approach Used To Evaluate Objective Quality Estimation Methods

Test Scene Selection. The selection of test scenes is a very important issue. In particular, the spatial and temporal information content of the scenes are critical parameters. These parameters play a crucial role in determining the amount of video compression that is possible, and consequently, the level of impairment that is suffered when the scene is transmitted over a fixed-rate digital transmission service channel. Figure A2 shows the relative amounts of spatial and temporal information for some possible test scenes. Fair and relevant video test scenes must be chosen such that their spatial and temporal information content is consistent with the video services that the digital transmission service channel was intended to provide. As shown in Figure A1 on page 26, two groups of test scenes are proposed; spatial test scenes to test the spatial performance, and temporal test scenes to test the temporal performance. The set of test scenes should span the full range of spatial and temporal information content of interest to users of this standard. The specific set of video scenes to use for testing should be agreed upon by the VTC/VT sub-working group as soon as possible.

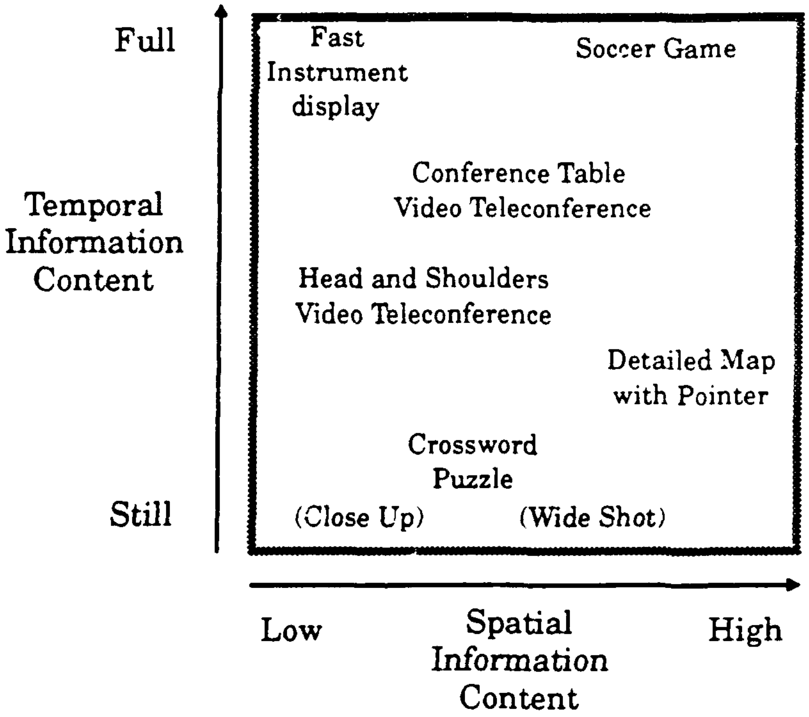


Figure A2 Information Content of Test Scenes

DUT Selection. The objective performance measures and algorithms set forth in this standard should be tested using a representative set of DUTs that include all relevant technologies for which this standard applies. Preferably, the pool of DUTs should be sufficiently large to test the technology independence (i.e., not dependent on the coding algorithm or transport architecture) of the objective quality estimators. Thus, selection of a DUT involves not only selection of the digital codec and bit rate,

but also selection of the transport mechanism. The specific set of DUTs to use for testing should be agreed upon by the VTC/VT sub-working group as soon as possible.

Subjective Viewing Tests. The subjective viewing tests should be conducted in accordance with CCIR Recommendation 500. The 5-point impairment scale given in Recommendation 500 is useful when comparison of subjective results between laboratories is desired, and when the range of impairments will vary considerably. The subject is first presented with a scene and then with an impaired version of the same scene. The subjects are instructed to decide on and mark the level of impairment in the second scene, using the first scene as a reference. The five possible responses are: Imperceptible (IP), Perceptible but Not Annoying (P/NA), Slightly Annoying (SA), Annoying (A), and Very Annoying (VA). This 5-point impairment scale intentionally covers a very wide range of impairment levels in a non-linear fashion. By including reference scenes, impairment tests take advantage of the fact that the human eye excels at making comparisons.

Subjective Combiners. The subjective combiners shown in Figure A1 on page 26 are required to combine the subjective impairment scores into an overall composite subjective score for the DUT. Figure A3 on page 30 gives a block diagram for the functions of the subjective combiners. For every DUT, the subjective test results produce a histogram of subjective impairment scores for each of the spatial test scenes ($s_1, s_2, ..., s_m$), and temporal test scenes ($t_1, t_2, ..., t_n$). These histograms are depicted graphically in Figure A3 by the heights of lines occurring in each of the histogram bins (IP, P/NA, SA, A, VA). It is important to note that, even for a fixed DUT, the histograms of subjective scores could vary dramatically depending upon the scene input. Some output scenes from the DUT may be rated IP while others may be rated VA. The purpose of the subjective combiners are to produce overall composite subjective scores for the spatial performance (S_s) and temporal performance (S_t) of the DUT from the histograms of subjective impairment scores. The first function that must be performed is to produce a subjective score for each scene ($S_{s1}, S_{s2}, ..., S_{sm}$, $S_{t1}, S_{t2}, ..., S_{tn}$). This is shown in Figure A3 as a histogram combiner. There are many possible ways in which the histograms could be collapsed. Three possible methods are:

1. The total percentage of responses where people were not annoyed (%IP+%P/NA). This might be a good indicator of customer satisfaction since people that are annoyed tend to complain.

2. The total percentage of responses where people do not see an impairment (%IP). This might be a good indicator of high quality video and hence might be useful for people that desire contribution quality video.

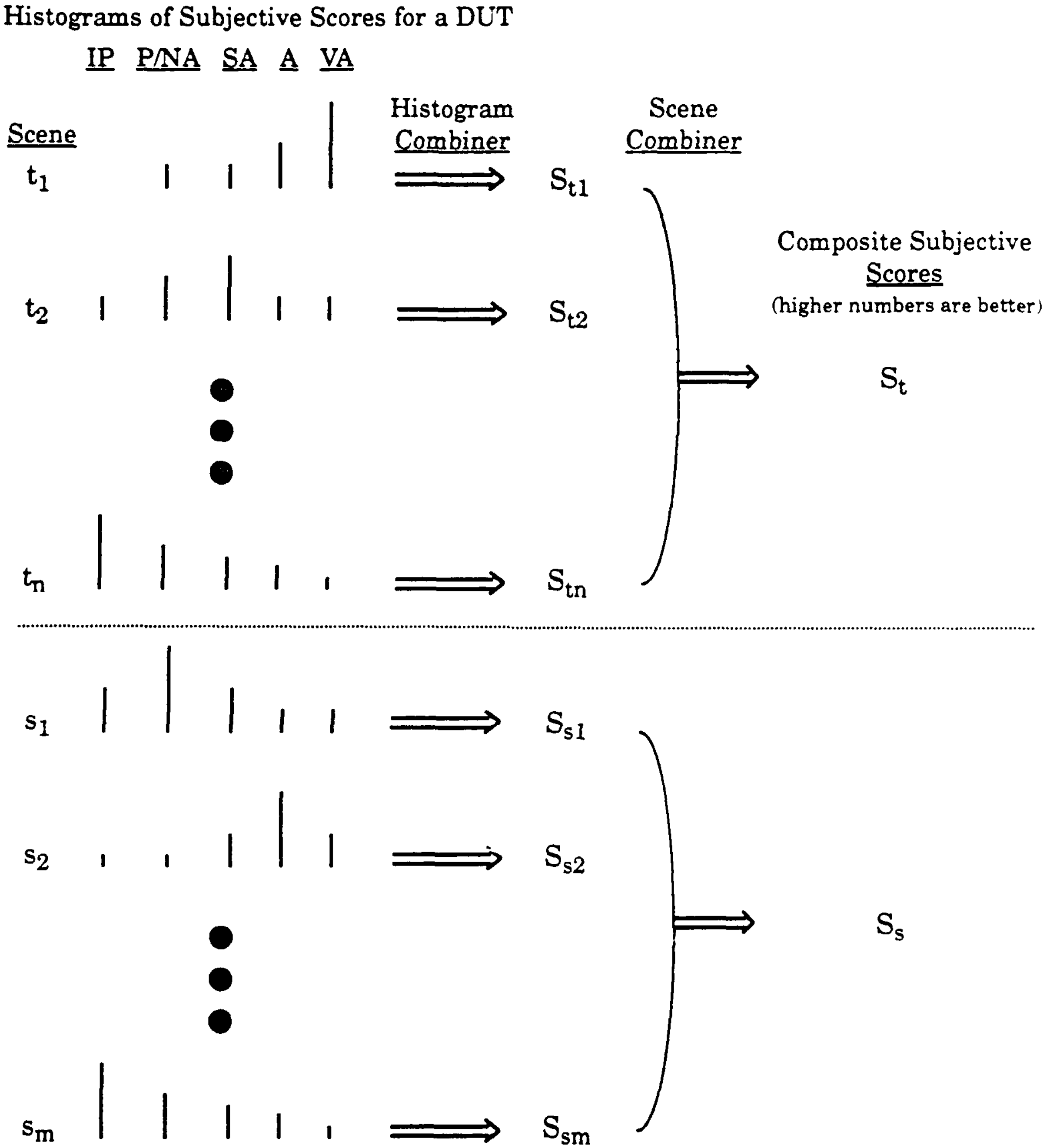
3. The weighted sum where each type of response is multiplied by a unique weight ($weight_1$(%VA) + $weight_2$(%A) + $weight_3$(%SA) + $weight_4$(%P/NA) + $weight_5$(%IP)). Here, the weight could vary in relationship to the importance of the particular histogram bin. A linear mean opinion score (MOS) can be generated from this equation if the weights are 1, 2, 3, 4, 5, respectively.

Next, the scene combiner shown in Figure A3 takes the subjective scores for each scene, as output by the histogram combiner (S_{s1}, S_{s2}, ..., S_{sm}, S_{t1}, S_{t2}, ..., S_{tn}), and produces the composite subjective scores for the spatial performance (S_s) and temporal performance (S_t) of the DUT. Once again, there are several possible methods to collapse the individual scene scores into overall composite scores for the DUT. Two possible methods are:

1. Assign equal weight (or importance) to each scene. Here, each spatial scene would have a weight of 1/m, and each temporal scene would have a weight of 1/n.

2. Assign zero weights to some of the test scenes depending upon the specific application as given in Table 1 on page 18. Then, assign equal weights to the remaining test scenes such that the sum of all spatial test scene weights is equal to one and the sum of all temporal test scene weights is equal to one. This method would allow the flexibility of tailoring the performance measurements to each of the applications given in Table 1 on page 18 (i.e., videophone - talking head, videophone - graphics, videoconference - people, etc.).

The specific subjective combiners should be agreed upon by the VTC/VT sub-working group as soon as possible.



NOTE:

The histograms of subjective scores for each DUT are a function of the scene input.

IP = Imperceptible
P/NA = Perceptible but Not Annoying
SA = Slightly Annoying
A = Annoying
VA = Very Annoying

Figure A3 Block Diagram of Subjective Combiners

Objective Test Waveforms. The objective test waveforms ($w_1, w_2, ..., w_j$) shown in Figure A1 on page 26 are analog video waveforms and include waveforms that are required to measure the traditional analog parameters as well as those that are required to measure any new objective parameters. After passing through the DUT, the output analog test waveforms are input into the proposed objective quality estimators (1 through k). Normally, a subset of the total waveforms will be utilized by each of the proposed objective quality estimators since some waveforms may have been specifically designed to work in conjunction with certain estimators.

Objective Quality Estimators. The proposed objective quality estimators (1 through k) shown in Figure A1 on page 26 each extract a set of objective performance parameters from the objective test waveforms. and combine these objective parameters to produce estimates of the spatial performance ($O_{s1}, O_{s2}, ..., O_{sk}$) and temporal performance ($O_{t1}, O_{t2}, ..., O_{tk}$) of the DUT. The designer of the objective quality estimator is free to choose the algorithm that produces quality estimates from the objective parameters. However, the criterion for determining the accuracy of the objective quality estimates is minimization of the averaged squared error difference between the objective and subjective estimates of performance over all DUTs. These squared error differences are denoted by E_s^2 and E_t^2 in Figure A1, where E_s is the error in the spatial performance estimate and E_t is the error in the temporal performance estimate for a DUT.

Objective Combiner. The objective combiner enables integration of the k objective quality estimates that are produced by the objective quality estimators. Composite estimates of spatial and temporal performance (O_s and O_t in Figure A1 on page 26) are produced by the objective combiner. The composite estimates will be at least as good as any of the individual estimates produced by the objective quality estimators. One could use linear predictors to predict O_s and O_t from the output of the objective quality estimators ($O_{s1}, O_{s2}, ..., O_{sk}$) and ($O_{t1}, O_{t2}, ..., O_{tk}$), respectively. Then, the prediction equations would take the form:

$$O_s = \left(\sum_{i=1}^{k} (\alpha_i \cdot O_{si})\right) + C_s$$

for the spatial performance and

$$O_t = \left(\sum_{i=1}^{k} (\beta_i \cdot O_{ti})\right) + C_t$$

for the temporal performance, where weights α_i and β_i are chosen such that the squared error differences between the objective and subjective estimates of performance are minimized over all DUTs. Constants C_s and C_t have been included to allow for the possibility that some of the individual estimates produced by the objective quality estimators are biased. One can look at the value of the individual weights α_i and β_i to determine if an objective quality estimator is not contributing to the overall composite score (the weights for that estimator will be small). In this case, the non-contributing estimator can be discarded without losing accuracy.

Test Results. This section presents detailed test results for the objective

quality measurement system used in this standard. The objective quality estimates obtained by use of this standard are compared with actual subjective quality ratings. Close correlation of the objective quality estimates with the subjective scores demonstrates the validity of the objective video quality estimators used in this standard. (Insert results of tests).

Glossary

Exchange Carrier (EC).

The telecommunications common carrier franchised to provide telecommunications services within one or more exchanges. An EC may also provide exchange access service, intra-LATA long-distance service, and in some unusual cases, inter-LATA service.

Interexchange Carrier (IC).

A telecommunications common carrier authorized to provide telecommunications services between LATAs. An IC may also provide service within some LATAs.

Local Access and Transport Area (LATA).

A geographic area established for the provision and administration of telecommunications services. A LATA encompasses one or more exchanges that have been grouped to serve common social, economic, and other purposes.

Network Interface (NI).

The point of demarcation between the carrier's facilities and the customer's installation which establishes the technical interface and division of operational responsibility.

Point of Termination (POT).

The point of demarcation between carriers which establishes the technical interface and division of operational responsibility.

Items Under Study